PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 9/10, 15/11, C08B 30/04

A1

(11) International Publication Number: WO 98/37213

(43) International Publication Date: 27 August 1998 (27.08.98)

(21) International Application Number: PCT/IB98/00270

(22) International Filing Date: 23 February 1998 (23.02.98)

(30) Priority Data:

9703663.6	21 February 1997 (21.02.97)	GB
9706060.2	24 March 1997 (24.03.97)	GB

(71) Applicant (*for all designated States except US*): DANISCO A/S [DK/DK]; Langebrogade 1, P.O. Box 17, DK-1001 Copenhagen K (DK).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): POULSEN, Peter [DK/DK]; Danisco a/s, Langebrogade 1, P.O. Box 17, DK-1001 Copenhagen K (DK).

(74) Agents: MASCHIO, Antonio et al.; D Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).

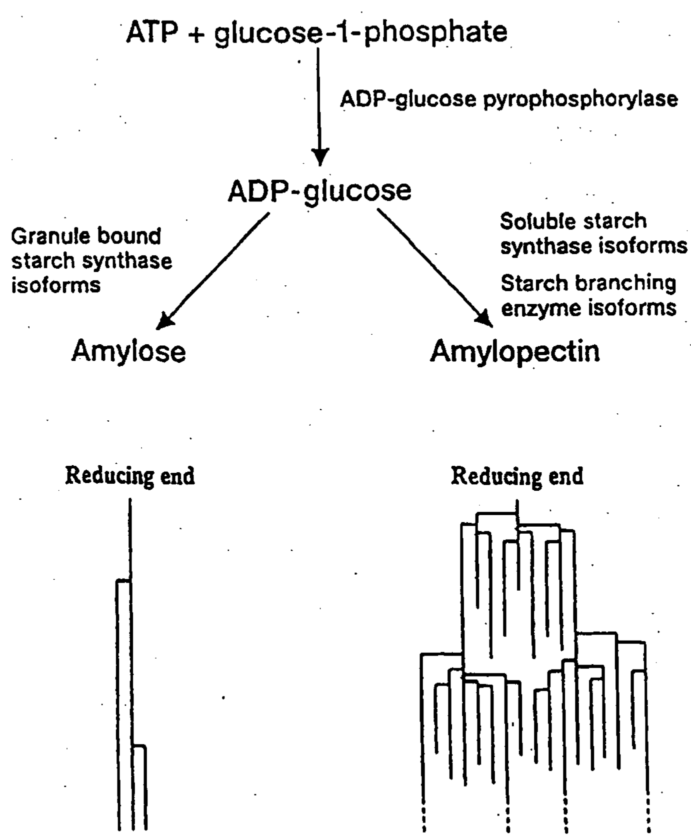
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: ANTISENSE INTRON INHIBITION OF STARCH BRANCHING ENZYME EXPRESSION

(57) Abstract

A method of inhibiting gene expression is described. The method, which affects enzymatic activity in a plant, comprises expressing in a plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron in an antisense orientation of a class A SBE; and wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

ANTISENSE INTRON INHIBITION OF STARCH BRANCHING ENZYME EXPRESSION

The present invention relates to a method of inhibiting gene expression, particularly inhibiting gene expression in a plant. The present invention also relates to a nucleotide sequence useful in the method. In addition, the present invention relates to a promoter that is useful for expressing the nucleotide sequence.

Starch is one of the main storage carbohydrates in plants, especially higher plants. The structure of starch consists of amylose and amylopectin. Amylose consists essentially of straight chains of α-1-4-linked glycosyl residues. Amylopectin comprises chains of α-1-4-linked glycosyl residues with some α-1-6 branches. The branched nature of amylopectin is accomplished by the action of *inter alia* an enzyme commonly known as the starch branching enzyme ("SBE"). SBE catalyses the formation of branch points in the amylopectin molecule by adding α-1,4 glucans through α-1,6-glucosidic branching linkages. The biosynthesis of amylose and amylopectin is schematically shown in Figure 1, whereas the α-1-4-links and the α-1-6 links are shown in Figure 2.

In Potato, it is known that two classes of SBE exist. In our copending international patent applications PCT/EP96/03052 and PCT/EP96/03053, class B potato SBE and a gene encoding it are discussed. In international patent application WO96/34968, class A potato SBE and a cDNA encoding it are disclosed.

It is known that starch is an important raw material. Starch is widely used in the food, paper, and chemical industries. However, a large fraction of the starches used in these industrial applications are post-harvest modified by chemical, physical or enzymatic methods in order to obtain starches with certain required functional properties.

Within the past few years it has become desirable to make genetically modified plants which could be capable of producing modified starches which could be the same as the post-harvest modified starches. It is also known that it may be possible to prepare such genetically modified plants by expression of antisense nucleotide coding sequences. In this regard, June Bourque provides a detailed summary of antisense strategies for the genetic manipulations in plants (Bourque 1995 Plant Science 105 pp 125-149). At this stage, reference could be made to Figure 3 which is a schematic diagram of one of the proposed mechanisms of antisense-RNA inhibition.

In particular, WO 92/11375 reports on a method of genetically modifying potato so as to form amylose-type starch. The method involves the use of an anti-sense construct that can apparently inhibit, to a varying extent, the expression of the gene coding for formation of the branching enzyme in potato. The antisense construct of WO 92/11375 consists of a tuber specific promoter, a transcription start sequence and the first exon of the branching enzyme in antisense direction. However, WO 92/11375 does not provide any antisense sequence data. In addition, WO 92/11375 only discloses the use of the potato GBSS promoter.

WO 92/14827 reports on a plasmid that, after insertion into the genome of a plant, can apparently cause changes in the carbohydrate concentration and carbohydrate composition, such as the concentration and composition of amylose and amylopectin, in the regenerated plant. The plasmid contains part of the coding sequence of a branching enzyme in an antisense orientation.

EP-A-0647715 reports on the use of antisense endogenous mRNA coding DNA to alter the characteristics and the metabolic pathways of ornamental plants.

EP-A-0467349 reports on the expression of sequences that are antisense to sequences upstream of a promoter to control gene expression.

EP-A-0458367 and US-A-5107065 report on the expression of a nucleotide sequence to regulate gene expression in a plant. The nucleotide sequence is complementary to a mRNA sequence of a gene and may cover all or a portion of the non-coding region of the gene. In other words, the nucleotide sequences of EP-A-0458367 and US-A-5107065 must at least comprise a sequence that is complementary to a coding region. EP-A-0458367 and US-A-5107065 contain minimal sequence information.

WO96/34968 discusses the use of antisense sequences complementary to sequences which encode class A and class B potato SBE to downregulate SBE expression in potato plants. The sequences used are complementary to SBE coding sequences.

Kuipers *et al* in Mol. Gen. Genet. [1995] 246 745-755 report on the expression of a series of nucleotides that are antisense to part of the genomic intron sequences of potato granule bound starch synthetase. Here the antisense intron sequences are attached to a part of the antisense exon sequences - wherein the intron sequences and the exon

sequences are naturally associated with each other. In addition, the expressed antisense intron sequences are at most 231 bp in length.

Likewise, Kull *et al* in J. Genet & Breed. [1995] 49 69-76 report on the expression of a series of nucleotides that are antisense to part of the genomic intron sequences of potato granule bound starch synthetase. Likewise, here the antisense intron sequences are attached to a part of the antisense exon sequences - wherein the intron sequences and the exon sequences are naturally associated with each other. In addition, likewise, the expressed antisense intron sequences are at most 231 bp in length.

Shimada *et al* in Theor. Appl. Genet. [1993] 86 665-672 report on the expression of a series of nucleotides that are antisense to part of the genomic intron sequences of rice granule bound starch synthetase. Here the antisense intron sequences are attached to a part of the antisense exon sequences - wherein the intron sequences and the exon sequences are naturally associated with each other. In addition, the expressed antisense intron sequences are less than 350 bp in length.

Reviews on how enzymatic activity can be affected by expression of particular nucleotide sequences may be found in the teachings of Finnegan and McElroy [1994] Biotechnology 12 883-888; and Matzke and Matzke [1995] TIG 11 1-3.

Whilst it is known that enzymatic activity can be affected by expression of particular nucleotide sequences there is still a need for a method that can more reliably and/or more efficiently and/or more specifically affect enzymatic activity.

According to a first aspect of the present invention there is provided a method of affecting enzymatic activity in a plant (or a cell, a tissue or an organ thereof) comprising expressing in the plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence partially or completely codes for (is) an intron of the potato class A SBE gene in an antisense orientation optionally together with a nucleotide sequence which codes, partially or completely, for an intron of a class B starch branching enzyme in an antisense or sense orientation; and wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

According to a second aspect of the present invention there is provided a method of affecting enzymatic activity in a starch producing organism (or a cell, a tissue or an

organ thereof) comprising expressing in the starch producing organism (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron of the potato class A SBE gene, in an antisense orientation optionally together with a nucleotide sequence which codes, partially or completely, for an intron of a class B starch branching enzyme in an antisense or sense orientation; and wherein starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed.

Preferably, the class A SBE gene antisense intron construct is used in combination with a potato class B SBE gene antisense intron construct as defined in PCT/EP96/03052. However, it may also be used independently thereof, to target class A SBE alone, or in combination with other transgenes, to further manipulate starch quality in potato plants.

According to a third aspect of the present invention, therefore, there is provided an antisense sequence comprising the nucleotide sequence shown as any one of SEQ.I.D. No. 15 to SEQ.I.D. No. 27 and the complement of SEQ. ID. No.38, or a variant, derivative or homologue thereof.

According to a fourth aspect of the present invention there is provided a promoter comprising the sequence shown as SEQ.I.D. No. 14 or a variant, derivative or homologue thereof.

According to a fifth aspect of the present invention there is provided a construct capable of comprising or expressing the present invention.

According to a sixth aspect of the present invention there is provided a vector comprising or expressing the present invention.

According to a seventh aspect of the present invention there is provided a cell, tissue or organ comprising or expressing the present invention.

According to an eighth aspect of the present invention there is provided a transgenic starch producing organism comprising or expressing the present invention.

According to a ninth aspect of the present invention there is provided a starch obtained from the present invention.

According to a tenth aspect of the present invention there is provided pSS17 and pSS18.

According to an eleventh aspect of the present invention there is provided a nucleotide sequence that is antisense to any one or more of the intron sequences obtainable from class A SBE, and especially those obtainable from intron 1of class A SBE as set forth in SEQ. ID. No. 38.

A key advantage of the present invention is that it provides a method for preparing modified starches that is not dependent on the need for post-harvest modification of starches. Thus the method of the present invention obviates the need for the use of hazardous chemicals that are normally used in the post-harvest modification of starches.

In addition, the present invention provides *inter alia* genetically modified plants which are capable of producing modified and/or novel and/or improved starches whose properties would satisfy various industrial requirements.

Thus, the present invention provides a method of preparing tailor-made starches in plants which could replace the post-harvest modified starches.

Also, the present invention provides a method that enables modified starches to be prepared by a method that can have a more beneficial effect on the environment than the known post-harvest modification methods which are dependent on the use of hazardous chemicals and large quantities of energy.

An other key advantage of the present invention is that it provides a method that may more reliably and/or more efficiently and/or more specifically affect enzymatic activity when compared to the known methods of affecting enzymatic activity. With regard to this advantage of the present invention it is to be noted that there is some degree of homology between coding regions of SBEs. However, there is little or no homology with the intron sequences of SBEs.

Thus, antisense intron expression provides a mechanism to affect selectively the expression of a particular class A SBE. This advantageous aspect could be used, for example, to reduce or eliminate a particular SBE enzyme, especially a class A SBE enzyme, and replace that enzyme with another enzyme which can be another branching enzyme or even a recombinant version of the affected enzyme or even a hybrid enzyme which could for example comprise part of a SBE enzyme from one source and at least a part of another SBE enzyme from another source. This particular feature of the present

invention is covered by the combination aspect of the present invention which is discussed in more detail later.

Thus the present invention provides a mechanism for selectively affecting class A SBE activity. This is in contrast to the prior art methods which are dependent on the use of for example antisense exon expression whereby it would not be possible to introduce new SBE activity without affecting that activity as well.

In the context of the present invention, class B SBE is synonymous with SBE I: class A SBE is synonymous with SBE II. Class A SBE is as defined in WO96/34968, incorporated herein by reference. Preferably, the antisense intron construct used comprises intron 1 of class A SBE, which is 2.0 kb in length and is located starting at residue 45 of the coding sequence of class A SBE. The boundaries of the intron may be calculated by searching for consensus intron boundary sequences, and are shown in attached figure 13. Class B SBE is substantially as defined in the sequences given herein and in PCT/EP96/03052.

Preferably with the first aspect of the present invention starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed.

Preferably with the second aspect of the present invention the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

Preferably with the fourth aspect of the present invention the promoter is in combination with a gene of interest ("GOI").

Preferably the enzymatic activity is reduced or eliminated.

Preferably the nucleotide sequence codes for at least substantially all of at least one intron in an antisense orientation.

Preferably the nucleotide sequence codes, partially or completely, for two or more introns and wherein each intron is in an anti-sense orientation.

Preferably the nucleotide sequence comprises at least 350 nucleotides (e.g. at least 350 bp), more preferably at least 500 nucleotides (e.g. at least 500 bp).

Preferably the nucleotide sequence comprises the complement of the sequence shown in SEQ. ID. No. 38, or a fragment thereof.

Preferably the nucleotide sequence is expressed by a promoter having a sequence shown as SEQ. I.D. No 14 or a variant, derivative or homologue thereof.

Preferably the transgenic starch producing organism is a plant.

A preferred aspect of the present invention therefore relates to a method of affecting enzymatic activity in a plant (or a cell, a tissue or an organ thereof) comprising expressing in the plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron in an antisense orientation; wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron; and wherein starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed.

A more preferred aspect of the present invention therefore relates to a method of affecting enzymatic activity in a plant (or a cell, a tissue or an organ thereof) comprising expressing in the plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron in an antisense orientation; wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron; wherein starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed; and wherein the nucleotide sequence comprises the sequence shown as any one of SEQ.I.D. No. 15 to SEQ.I.D. No. 27 or a variant, derivative or homologue thereof, including combinations thereof.

The term "nucleotide" in relation to the present invention includes DNA and RNA. Preferably it means DNA, more preferably DNA prepared by use of recombinant DNA techniques.

The term "intron" is used in its normal sense as meaning a segment of nucleotides, usually DNA, that is transcribed but does not encode part or all of an expressed protein or enzyme.

The term "exon" is used in its normal sense as meaning a segment of nucleotides, usually DNA, encoding part or all of an expressed protein or enzyme.

Thus, the term "intron" refers to gene regions that are transcribed into RNA molecules, but which are spliced out of the RNA before the RNA is translated into a

protein. In contrast, the term "exon" refers to gene regions that are transcribed into RNA and subsequently translated into proteins.

The terms "variant" or "homologue" or "fragment" in relation to the nucleotide sequence of the present invention include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the respective nucleotide sequence providing the resultant nucleotide sequence can affect enzyme activity in a plant, or cell or tissue thereof, preferably wherein the resultant nucleotide sequence has at least the same effect as the complement of the sequence shown as SEQ.I.D. No. 38. In particular, the term "homologue" covers homology with respect to similarity of structure and/or similarity of function providing the resultant nucleotide sequence has the ability to affect enzymatic activity in accordance with the present invention. With respect to sequence homology (i.e. similarity), preferably there is more than 80% homology, more preferably at least 85% homology, more preferably at least 90% homology, even more preferably at least 95% homology, more preferably at least 98% homology. The above terms are also synonymous with allelic variations of the sequences.

Likewise, the terms "variant" or "homologue" or "fragment" in relation to the promoter of the present invention include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the respective promoter sequence providing the resultant promoter sequence allows expression of a GOI, preferably wherein the resultant promoter sequence has at least the same effect as SEQ.I.D. No. 14. In particular, the term "homologue" covers homology with respect to similarity of structure and/or similarity of function providing the resultant promoter sequence has the ability to allow for expression of a GOI, such as a nucleotide sequence according to the present invention. With respect to sequence homology (i.e. similarity), preferably there is more than 80% homology, more preferably at least 85% homology, more preferably at least 90% homology, even more preferably at least 95% homology, more preferably at least 98% homology. The above terms are also synonymous with allelic variations of the sequences.

The term "antisense" means a nucleotide sequence that is complementary to, and can therefore hybridise with, any one or all of the intron sequences of the present invention, including partial sequences thereof.

With the present invention, the antisense intron can be complementary to an entire intron of the gene to be inhibited. However, in some circumstances, partial antisense sequences may be used (i.e. sequences that are not or do not comprise the full complementary sequence) providing the partial sequences affect enzymatic activity. Suitable examples of partial sequences include sequences that are shorter than the full complement of SEQ. ID. No. 38 but which comprise nucleotides that are at least antisense to the sense intron sequences adjacent the respective exon or exons.

With regard to the second aspect of the present invention (i.e. specifically affecting SBE activity), the nucleotide sequences of the present invention may comprise one or more sense or antisense exon sequences of the SBE gene, including complete or partial sequences thereof, providing the nucleotide sequences can affect SBE activity, preferably wherein the nucleotide sequences reduce or eliminate SBE activity. Preferably, the nucleotide sequence of the second aspect of the present invention does not comprise an antisense exon sequence.

The term "vector" includes an expression vector and a transformation vector. The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression. The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E. Coli* plasmid to a fungus or a plant cell, or from an *Agrobacterium* to a plant cell.

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - in relation to the antisense nucleotide sequence aspect of the present invention includes the nucleotide sequence according to the present invention directly or indirectly attached to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the *Sh1*-intron or the ADH intron, intermediate the promoter and the nucleotide sequence of the present invention. The same is true for the term "fused" in relation to the present invention which includes direct or indirect attachment. The terms do not cover the natural

combination of the wild type SBE gene when associated with the wild type SBE gene promoter in their natural environment.

The construct may even contain or express a marker which allows for the selection of the genetic construct in, for example, a plant cell into which it has been transferred. Various markers exist which may be used in, for example, plants - such as mannose. Other examples of markers include those that provide for antibiotic resistance - e.g. resistance to G418, hygromycin, bleomycin, kanamycin and gentamycin.

The construct of the present invention preferably comprises a promoter. The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Monod theory of gene expression. Examples of suitable promoters are those that can direct efficient expression of the nucleotide sequence of the present invention and/or in a specific type of cell. Some examples of tissue specific promoters are disclosed in WO 92/11375.

The promoter could additionally include conserved regions such as a Pribnow Box or a TATA box. The promoters may even contain other sequences to affect (such as to maintain, enhance, decrease) the levels of expression of the nucleotide sequence of the present invention. Suitable examples of such sequences include the *Sh1*-intron or an ADH intron. Other sequences include inducible elements - such as temperature, chemical, light or stress inducible elements. Also, suitable elements to enhance transcription or translation may be present. An example of the latter element is the TMV 5' leader sequence (see Sleat Gene 217 [1987] 217-225; and Dawson Plant Mol. Biol. 23 [1993] 97).

As mentioned, the construct and/or the vector of the present invention may include a transcriptional initiation region which may provide for regulated or constitutive expression. Any suitable promoter may be used for the transcriptional initiation region, such as a tissue specific promoter. In one aspect, preferably the promoter is the patatin promoter or the E35S promoter. In another aspect, preferably the promoter is the SBE promoter.

If, for example, the organism is a plant then the promoter can be one that affects expression of the nucleotide sequence in any one or more of seed, tuber, stem, sprout, root and leaf tissues, preferably tuber. By way of example, the promoter for the

nucleotide sequence of the present invention can be the α-Amy 1 promoter (otherwise known as the Amy 1 promoter, the Amy 637 promoter or the α-Amy 637 promoter) as described in our co-pending UK patent application No. 9421292.5 filed 21 October 1994. Alternatively, the promoter for the nucleotide sequence of the present invention can be the α-Amy 3 promoter (otherwise known as the Amy 3 promoter, the Amy 351 promoter or the α-Amy 351 promoter) as described in our co-pending UK patent application No. 9421286.7 filed 21 October 1994.

The present invention also encompasses the use of a promoter to express a nucleotide sequence according to the present invention, wherein a part of the promoter is inactivated but wherein the promoter can still function as a promoter. Partial inactivation of a promoter in some instances is advantageous. In particular, with the Amy 351 promoter mentioned earlier it is possible to inactivate a part of it so that the partially inactivated promoter expresses the nucleotide sequence of the present invention in a more specific manner such as in just one specific tissue type or organ. The term "inactivated" means partial inactivation in the sense that the expression pattern of the promoter is modified but wherein the partially inactivated promoter still functions as a promoter. However, as mentioned above, the modified promoter is capable of expressing a gene coding for the enzyme of the present invention in at least one (but not all) specific tissue of the original promoter. Examples of partial inactivation include altering the folding pattern of the promoter sequence, or binding species to parts of the nucleotide sequence, so that a part of the nucleotide sequence is not recognised by, for example, RNA polymerase. Another, and preferable, way of partially inactivating the promoter is to truncate it to form fragments thereof. Another way would be to mutate at least a part of the sequence so that the RNA polymerase can not bind to that part or another part. Another modification is to mutate the binding sites for regulatory proteins for example the CreA protein known from filamentous fungi to exert carbon catabolite repression, and thus abolish the catabolite repression of the native promoter.

The construct and/or the vector of the present invention may include a transcriptional termination region.

The nucleotide according to the present invention can be expressed in combination (but not necessarily at the same time) with an additional construct. Thus the present

invention also provides a combination of constructs comprising a first construct comprising the nucleotide sequence according to the present invention operatively linked to a first promoter; and a second construct comprising a GOI operatively linked to a second promoter (which need not be the same as the first promoter). With this aspect of the present invention the combination of constructs may be present in the same vector, plasmid, cells, tissue, organ or organism. This aspect of the present invention also covers methods of expressing the same, preferably in specific cells or tissues, such as expression in just a specific cell or tissue, of an organism, typically a plant. With this aspect of the present invention the second construct does not cover the natural combination of the gene coding for an enzyme ordinarily associated with the wild type gene promoter when they are both in their natural environment.

An example of a suitable combination would be a first construct comprising the nucleotide sequence of the present invention and a promoter, such as the promoter of the present invention, and a second construct comprising a promoter, such as the promoter of the present invention, and a GOI wherein the GOI codes for another starch branching enzyme either in sense or antisense orientation.

The above comments relating to the term "construct" for the antisense nucleotide aspect of the present invention are equally applicable to the term "construct" for the promoter aspect of the present invention. In this regard, the term includes the promoter according to the present invention directly or indirectly attached to a GOI.

The term "GOI" with reference to the promoter aspect of the present invention or the combination aspect of the present invention means any gene of interest, which need not necessarily code for a protein or an enzyme - as is explained later. A GOI can be any nucleotide sequence that is either foreign or natural to the organism in question, for example a plant.

Typical examples of a GOI include genes encoding for other proteins or enzymes that modify metabolic and catabolic processes. The GOI may code for an agent for introducing or increasing pathogen resistance.

The GOI may even be an antisense construct for modifying the expression of natural transcripts present in the relevant tissues. An example of such a GOI is the nucleotide sequence according to the present invention.

The GOI may even code for a protein that is non-natural to the host organism - e.g. a plant. The GOI may code for a compound that is of benefit to animals or humans. For example, the GOI could code for a pharmaceutically active protein or enzyme such as any one of the therapeutic compounds insulin, interferon, human serum albumin, human growth factor and blood clotting factors. The GOI may even code for a protein giving additional nutritional value to a food or feed or crop. Typical examples include plant proteins that can inhibit the formation of anti-nutritive factors and plant proteins that have a more desirable amino acid composition (e.g. a higher lysine content than a non-transgenic plant). The GOI may even code for an enzyme that can be used in food processing such as xylanases and α-galactosidase. The GOI can be a gene encoding for any one of a pest toxin, an antisense transcript such as that for α-amylase, a protease or a glucanase. Alternatively, the GOI can be a nucleotide sequence according to the present invention.

The GOI can be the nucleotide sequence coding for the arabinofuranosidase enzyme which is the subject of our co-pending UK patent application 9505479.7. The GOI can be the nucleotide sequence coding for the glucanase enzyme which is the subject of our co-pending UK patent application 9505475.5. The GOI can be the nucleotide sequence coding for the α-amylase enzyme which is the subject of our co-pending UK patent application 9413439.2. The GOI can be the nucleotide sequence coding for the α-amylase enzyme which is the subject of our co-pending UK patent application 9421290.9. The GOI can be any of the nucleotide sequences coding for the α-glucan lyase enzyme which are described in our co-pending PCT patent application PCT/EP94/03397.

In one aspect the GOI can even be a nucleotide sequence according to the present invention but when operatively linked to a different promoter.

The GOI could include a sequence that codes for one or more of a xylanase, an arabinase, an acetyl esterase, a rhamnogalacturonase, a glucanase, a pectinase, a branching enzyme or another carbohydrate modifying enzyme or proteinase. Alternatively, the GOI may be a sequence that is antisense to any of those sequences.

As mentioned above, the present invention provides a mechanism for selectively affecting a particular enzymatic activity. In an important application of the present invention it is now possible to reduce or eliminate expression of a genomic nucleotide sequence coding for a genomic protein or enzyme by expressing an antisense intron construct for that particular genomic protein or enzyme and (e.g. at the same time) expressing a recombinant version of that enzyme or protein - in other words the GOI is a recombinant nucleotide sequence coding for the genomic enzyme or protein. This application allows expression of desired recombinant enzymes and proteins in the absence of (or reduced levels of) respective genomic enzymes and proteins. Thus the desired recombinant enzymes and proteins can be easily separated and purified from the host organism. This particular aspect of the present invention is very advantageous over the prior art methods which, for example, rely on the use of anti-sense exon expression which methods also affect expression of the recombinant enzyme.

Thus, a further aspect of the present invention relates to a method of expressing a recombinant protein or enzyme in a host organism comprising expressing a nucleotide sequence coding for the recombinant protein or enzyme; and expressing a further nucleotide sequence wherein the further nucleotide sequence codes, partially or completely, for an intron in an antisense orientation; wherein the intron is an intron normally associated with the genomic gene encoding a protein or an enzyme corresponding to the recombinant protein or enzyme; and wherein the further nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron. Additional aspects cover the combination of those nucleotide sequences including their incorporation in constructs, vectors, cells, tissues and transgenic organisms.

Therefore the present invention also relates to a combination of nucleotide sequences comprising a first nucleotide sequence coding for a recombinant enzyme; and a second nucleotide sequence which corresponds to an intron in antisense orientation; wherein the intron is an intron that is associated with a genomic gene encoding an enzyme corresponding to the recombinant enzyme; and wherein the second nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

The GOI may even code for one or more introns, such as any one or more of the intron sequences presented in the attached sequence listings. For example, the present invention also covers the expression of for example an antisense intron (e.g. the complement of SEQ. ID. No. 38) in combination with for example a sense intron which preferably is not complementary to the antisense intron sequence (e.g. SEQ.I.D.No. 2 or another class A SBE intron).

The terms "cell", "tissue" and "organ" include cell, tissue and organ *per se* and when within an organism.

The term "organism" in relation to the present invention includes any organism that could comprise the nucleotide sequence according to the present invention and/or wherein the nucleotide sequence according to the present invention can be expressed when present in the organism. Preferably the organism is a starch producing organism such as any one of a plant, algae, fungi, yeast and bacteria, as well as cell lines thereof. Preferably the organism is a plant.

The term "starch producing organism" includes any organism that can biosynthesise starch. Preferably, the starch producing organism is a plant.

The term "plant" as used herein includes any suitable angiosperm, gymnosperm, monocotyledon and dicotyledon. Typical examples of suitable plants include vegetables such as potatoes; cereals such as wheat, maize, and barley; fruit; trees; flowers; and other plant crops. Preferably, the term means "potato".

The term "transgenic organism" in relation to the present invention includes any organism that comprises the nucleotide sequence according to the present invention and/or products obtained therefrom, and/or wherein the nucleotide sequence according to the present invention can be expressed within the organism. Preferably the nucleotide sequence of the present invention is incorporated in the genome of the organism. Preferably the transgenic organism is a plant, more preferably a potato.

To prepare the host organism one can use prokaryotic or eukaryotic organisms. Examples of suitable prokaryotic hosts include *E. coli* and *Bacillus subtilis*. Teachings on the transformation of prokaryotic hosts is well documented in the art, for example see Sambrook *et al* (Sambrook *et al.* in Molecular Cloning: A Laboratory Manual, 2nd edition. 1989, Cold Spring Harbor Laboratory Press).

Even though the enzyme according to the present invention and the nucleotide sequence coding for same are not disclosed in EP-B-0470145 and CA-A-2006454, those two documents do provide some useful background commentary on the types of techniques that may be employed to prepare transgenic plants according to the present invention. Some of these background teachings are now included in the following commentary.

The basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted genetic material.

Several techniques exist for inserting the genetic information, the two main principles being direct introduction of the genetic information and introduction of the genetic information by use of a vector system. A review of the general techniques may be found in articles by Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

Thus, in one aspect, the present invention relates to a vector system which carries a nucleotide sequence or construct according to the present invention and which is capable of introducing the nucleotide sequence or construct into the genome of an organism, such as a plant.

The vector system may comprise one vector, but it can comprise two vectors. In the case of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung An *et al.* (1980), Binary Vectors, *Plant Molecular Biology Manual A3*, 1-19.

One extensively employed system for transformation of plant cells with a given promoter or nucleotide sequence or construct is based on the use of a Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid from *Agrobacterium rhizogenes* An *et al.* (1986), *Plant Physiol.* 81, 301-305 and Butcher D.N. *et al.* (1980), *Tissue Culture Methods for Plant Pathologists,* eds.: D.S. Ingrams and J.P. Helgeson, 203-208. Several different Ti and Ri plasmids have been constructed which are suitable for the construction of the plant or plant cell constructs described above. A non-limiting example of such a Ti plasmid is pGV3850.

The nucleotide sequence or construct of the present invention should preferably be inserted into the Ti-plasmid between the terminal sequences of the T-DNA or adjacent a T-DNA sequence so as to avoid disruption of the sequences immediately surrounding the T-DNA borders, as at least one of these regions appears to be essential for insertion of modified T-DNA into the plant genome.

As will be understood from the above explanation, if the organism is a plant the vector system of the present invention is preferably one which contains the sequences necessary to infect the plant (e.g. the *vir* region) and at least one border part of a T-DNA sequence, the border part being located on the same vector as the genetic construct.

Furthermore, the vector system is preferably an *Agrobacterium tumefaciens* Ti-plasmid or an *Agrobacterium rhizogenes* Ri-plasmid or a derivative thereof. As these plasmids are well-known and widely employed in the construction of transgenic plants, many vector systems exist which are based on these plasmids or derivatives thereof.

In the construction of a transgenic plant the nucleotide sequence or construct of the present invention may be first constructed in a microorganism in which the vector can replicate and which is easy to manipulate before insertion into the plant. An example of a useful microorganism is *E. coli*, but other microorganisms having the above properties may be used. When a vector of a vector system as defined above has been constructed in *E. coli*, it is transferred, if necessary, into a suitable *Agrobacterium* strain, e.g. *Agrobacterium tumefaciens*. The Ti-plasmid harbouring the nucleotide sequence or construct of the present invention is thus preferably transferred into a suitable *Agrobacterium* strain, e.g. *A. tumefaciens*, so as to obtain an *Agrobacterium* cell harbouring the promoter or nucleotide sequence or construct of the present invention, which DNA is subsequently transferred into the plant cell to be modified.

If, for example, for the transformation the Ti- or Ri-plasmid of the plant cells is used, at least the right boundary and often however the right and the left boundary of the Ti- and Ri-plasmid T-DNA, as flanking areas of the introduced genes, can be connected. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema, in: The Binary Plant Vector System Offset-drukkerij Kanters B.B., Alblasserdam, 1985, Chapter V; Fraley, *et al.*, Crit. Rev. Plant Sci., 4:1-46; and An *et al.*, EMBO J. (1985) 4:277-284.

Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher D.N. *et al.* (1980), *Tissue Culture Methods for Plant Pathologists,* eds.: D.S. Ingrams and J.P. Helgeson, 203-208. For further teachings on this topic see Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27). With this technique, infection of a plant may be performed in or on a certain part or tissue of the plant, i.e. on a part of a leaf, a root, a stem or another part of the plant.

Typically, with direct infection of plant tissues by *Agrobacterium* carrying the GOI (such as the nucleotide sequence according to the present invention) and, optionally, a promoter, a plant to be infected is wounded, e.g. by cutting the plant with a razor blade or puncturing the plant with a needle or rubbing the plant with an abrasive. The wound is then inoculated with the *Agrobacterium*. The inoculated plant or plant part is then grown on a suitable culture medium and allowed to develop into mature plants.

When plant cells are constructed, these cells may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

Regeneration of the transformed cells into genetically modified plants may be accomplished using known methods for the regeneration of plants from cell or tissue cultures, for example by selecting transformed shoots using an antibiotic and by subculturing the shoots on a medium containing the appropriate nutrients, plant hormones, etc.

Further teachings on plant transformation may be found in EP-A-0449375.

As reported in CA-A-2006454, a large amount of cloning vectors are available which contain a replication system in *E. coli* and a marker which allows a selection of the transformed cells. The vectors contain for example pBR 322, pUC series, M13 mp series, pACYC 184 etc. In this way, the nucleotide or construct of the present invention can be introduced into a suitable restriction position in the vector. The contained plasmid is then used for the transformation in *E.coli*. The *E.coli* cells are cultivated in a suitable nutrient medium and then harvested and lysed. The plasmid is then recovered. As a method of analysis there is generally used sequence analysis, restriction analysis,

electrophoresis and further biochemical-molecular biological methods. After each manipulation, the used DNA sequence can be restricted and connected with the next DNA sequence. Each sequence can be cloned in the same or different plasmid.

After the introduction of the nucleotide sequence or construct according to the present invention in the plants the presence and/or insertion of further DNA sequences may be necessary - such as to create combination systems as outlined above (e.g. an organism comprising a combination of constructs).

The above commentary for the transformation of prokaryotic organisms and plants with the nucleotide sequence of the present invention is equally applicable for the transformation of those organisms with the promoter of the present invention.

In summation, the present invention relates to affecting enzyme activity by expressing antisense intron sequences.

Also, the present invention relates to a promoter useful for the expression of those antisense intron sequences.

The following samples have been deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St Machar Drive, Aberdeen, Scotland, AB2 1RY, United Kingdom, on 13 July 1995:

NCIMB 40753 (which refers to pBEA 8 as described herein);

NCIMB 40751 (which refers to λ-SBE 3.2 as described herein), and

NCIMB 40752 (which refers to λ-SBE 3.4 as described herein).

The following sample has been deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St Machar Drive, Aberdeen, Scotland, AB2 1RY, United Kingdom, on 9 July 1996:

NCIMB 40815 (which refers to pBEA 9 as described herein).

A highly preferred embodiment of the present invention therefore relates to a method of affecting enzymatic activity in a plant (or a cell, a tissue or an organ thereof) comprising expressing in the plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron in an antisense orientation; wherein the nucleotide sequence does not contain a sequence that

is antisense to an exon sequence normally associated with the intron; wherein starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed; and wherein the nucleotide sequence is antisense to intron 1 of class A SBE as set forth in SEQ. ID. No. 38, or any other intron of class A SBE, including fragments thereof, and including combinations of class A antisense intron sequences and class B antisense intron sequences. The sequence of introns of class A SBE other than intron 1 may be obtained by sequencing of, for example, potato class A SBE genomic DNA, isolatable by hybridisation screening of a genomic DNA library with class A SBE cDNA obtainable according to WO96/34968 according to methods well known in the art and set forth, for example, in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989.

The present invention will now be described only by way of example, in which reference is made to the following attached Figures:

Figure 1, which is a schematic representation of the biosynthesis of amylose and amylopectin;

Figure 2, which is a diagrammatic representation of the α-1-4-links and the α-1-6 links of amylopectin;

Figure 3, which is a diagrammatic representation of a possible antisense-RNA inhibition mechanism;

Figure 4, which is a diagrammatic representation of the exon-intron structure of a genomic SBE clone;

Figure 5, which is a plasmid map of pPATA1, which is 3936 bp in size;

Figure 6, which is a plasmid map of pABE6, which is 5106 bp in size;

Figure 7, which is a plasmid map of pVictorIV Man, which is 7080 bp in size;

Figure 8, which is a plasmid map of pBEA8, which is 9.54 kb in size;

Figure 9, which is a plasmid map of pBEA9, which is 9.54 kb in size;

Figure 10, which is a plasmid map of pBEP2, which is 10.32 kb in size;

Figure 11, which is a plasmid map of pVictor5a, which is 9.12 kb in size;

Figure 12, which shows the full genomic nucleotide sequence for SBE including the promoter, exons and introns;

Figure 13, which shows the positioning of intron 1 in the class A and class B SBE genes;

Figure 14, which shows the sequence of intron 1 of the potato class A SBE;

Figure 15, which shows the structure of pSS17; and

Figure 16, which shows the structure of pSS18.

Figures 1 and 2 were referred to above in the introductory description concerning starch in general. Figure 3 was referred to above in the introductory description concerning antisense expression.

As mentioned, Figure 4 is a diagrammatic representation of the exon-intron structure of a genomic SBE clone, the sequence of which is shown in Figure 12. This clone, which has about 11.5 k base pairs, comprises 14 exons and 13 introns. The introns are numbered in increasing order from the 5' end to the 3' end and correspond to SEQ.I.D.No.s 1-13, respectively. Their respective antisense intron sequences are shown as SEQ.I.D.No.s 15-27.

In more detail, Figures 4 and 12 present information on the 11478 base pairs of a potato SBE gene. The 5' region from nucleotides 1 to 2082 contain the promoter region of the SBE gene. A TATA box candidate at nucleotide 2048 to 2051 is boxed. The homology between a potato SBE cDNA clone (Poulsen & Kreiberg (1993) Plant Physiol 102: 1053-1054) and the exon DNAs begin at 2083 bp and end at 9666 bp.

The homology between the cDNA and the exon DNA is indicated by nucleotides in upper case letters, while the translated amino acid sequences are shown in the single letter code below the exon DNA. Intron sequences are indicated by lower case letters.

Figures 5 to 7 are discussed below. As mentioned, Figure 8 is a plasmid map of pBEA8, which is 9.54 k base pairs in size; and Figure 9 is a plasmid map of pBEA9, which is 9.54 k base pairs in size. Each of pBEA 8 and pBEA 9 comprises an antisense sequence to the first intron sequence of the potato SBE gene. This first intron sequence, which has 1177 base pairs, is shown in Figure 4 and lies between the first exon and the second exon.

These experiments and aspects of the present invention are now discussed in more detail.

EXPERIMENTAL PROTOCOL

ISOLATION, SUBCLONING IN PLASMIDS, AND SEQUENCING OF GENOMIC CLASS B SBE CLONES

Various clones containing the potato class B SBE gene are isolated from a Desiree potato genomic library (Clontech Laboratories Inc., Palo Alto CA, USA) using radioactively labelled potato SBE cDNA (Poulsen & Kreiberg (1993) Plant Physiol. 102:1053-1054) as probe. The fragments of the isolated λ-phages containing SBE DNA (λSBE 3.2 - NCIMB 40751 - and λSBE-3.4 - NCIMB 40752) are identified by Southern analysis and then subcloned into pBluescript II vectors (Clontech Laboratories Inc., Palo Alto CA, USA). λSBE 3.2 contains a 15 kb potato DNA insert and λSBE-3.4 contains a 13 kb potato DNA insert. The resultant plasmids are called pGB3, pGB11, pGB15, pGB16 and pGB25 (see discussion below). The respective inserts are then sequenced using the Pharmacia Autoread Sequencing Kit (Pharmacia, Uppsala) and a A.L.F. DNA sequencer (Pharmacia, Uppsala).

In total, a stretch of 11.5 kb of the class B SBE gene is sequenced. The sequence is deduced from the above-mentioned plasmids, wherein: pGB25 contains the sequences from 1 bp to 836 bp, pGB15 contains the sequences from 735 bp to 2580 bp, pGB16 contains the sequences from 2580 bp to 5093 bp, pGB11 contains the sequences from 3348 bp to 7975 bp, and pGB3 contains the sequences from 7533 bp to 11468 bp.

In more detail, pGB3 is constructed by insertion of a 4 kb *Eco*RI fragment isolated from λSBE 3.2 into the *Eco*RI site of pBluescript II SK (+). pGB11 is constructed by insertion of a 4.7 kb *Xho*I fragment isolated from λSBE 3.4 into the *Xho*I site of pBluescript II SK (+). pGB15 is constructed by insertion of a 1.7 kb *Spe*I fragment isolated from λSBE 3.4 into the *Spe*I site of pBluescript II SK (+). pGB16 is constructed by insertion of a 2.5 kb *Spe*I fragment isolated from λSBE 3.4 into the *Spe*I site of pBluescript II SK (+). For the construction of pGB25 a PCR fragment is produced with the primers

5' GGA ATT CCA GTC GCA GTC TAC ATT AC 3'

(SEQ. ID. No.30)

and

5' CGG GAT CCA GAG GCA TTA AGA TTT CTG G 3'

(SEQ. ID. No. 31)

and λSBE 3.4 as a template.

The PCR fragment is digested with *BamHI* and *EcoRI*, and inserted in pBluescript II SK (+) digested with the same restriction enzymes.

A class A SBE clone is derived similarly.

CONSTRUCTION OF CLASS B SBE ANTISENSE INTRON PLASMIDS pBEA8 and pBEA9

The SBE intron 1 is amplified by PCR using the oligonucleotides:

5' CGG GAT CCA AAG AAA TTC TCG AGG TTA CAT GG 3'

(SEQ. ID. No. 32)

and

5' CGG GAT CCG GGG TAA TTT TTA CTA ATT TCA TG 3'

(SEQ. ID. No. 33)

and the λSBE 3.4 phage containing the SBE gene as template.

The PCR product is digested with *BamHI* and inserted in an antisense orientation in the *BamHI* site of plasmid pPATA1 (described in WO 94/24292) between the patatin promoter and the 35S terminator. This construction, pABE6, is digested with *KpnI*, and the 2.4 kb "patatin promoter-SBE intron 1- 35S terminator" *KpnI* fragment is isolated and inserted in the *KpnI* site of the plant transformation vector pVictorIV Man. The *KpnI* fragment is inserted in two orientations yielding plasmids pBEA8 and pBEA9. pVictorIV Man is shown in Figure 7 and is formed by insertion of a filled in *XbaI* fragment containing a E35S promoter-manA-35S terminator cassette isolated from plasmid pVictorIV SGiN Man (WO 94/24292) into the filled in *XhoI* site of pVictor IV. The pVictor regions of pVictor IV Man contained between the co-ordinates 2.52 bp to 0.32 bp (see Figure 7).

CONSTRUCTION OF CLASS A SBE ANTISENSE INTRON PLASMIDS pSS17 and pSS18

Construction of plasmid pSS17.

The 2122 bp intron 1 sequence of the potato SBEII gene is amplified by PCR from a genomic SBEII subclone using the primers 5' - CGG GAT CCC GTA TGT CTC ACT GTG TTT GTG GC - 3' (SEQ. ID. No. 34) and 5' - CGG GAT CCC CCT ACA TAC ATA TAT CAG ATT AG - 3' (SEQ. ID. No. 35). The PCR product is digested with BamHI and inserted in antisense orientation after a patatin promoter in the BamHI site of a plant transformation vector in which the NPTII gene is used as selectable marker (see figure 15).

Construction of plasmid pSS18.

The 2122 bp intron 1 sequence of the potato SBEII gene is amplified by PCR from a genomic SBEII subclone using the primers 5' - CGG GAT CCC GTA TGT CTC ACT GTG TTT GTG GC - 3' (SEQ. ID. No. 34) and 5' - CGG GAT CCC CCT ACA TAC ATA TAT CAG ATT AG - 3' (SEQ. ID. No. 35). The PCR product is digested with BamHI and inserted in antisense orientation after a patatin promoter in the BamHI site of a plant transformation vector in which the *manA* gene is used as selectable marker (see figure 16).

PRODUCTION OF TRANSGENIC POTATO PLANTS

<u>Axenic stock cultures</u>

Shoot cultures of *Solanum tuberosum* 'Bintje' and 'Dianella' are maintained on a substrate (LS) of a formula according to Linsmaier, E.U. and Skoog, F. (1965), Physiol. Plant. 18: 100-127, in addition containing 2 μM silver thiosulphate at 25°C and 16 h light/8 h dark.

The cultures are subcultured after approximately 40 days. Leaves are then cut off the shoots and cut into nodal segments (approximately 0.8 cm) each containing one node.

Inoculation of potato tissues

Shoots from approximately 40 days old shoot cultures (height approximately 5-6 cms) are cut into internodal segments (approximately 0.8 cm). The segments are placed into liquid LS-substrate containing the transformed *Agrobacterium tumefaciens* containing the binary vector of interest. The *Agrobacterium* are grown overnight in YMB-substrate (di-potassium hydrogen phosphate, trihydrate (0.66 g/l); magnesium sulphate, heptahydrate (0.20 g/l); sodium chloride (0.10 g/l); mannitol (10.0 g/l); and yeast extract (0.40 g/l)) containing appropriate antibiotics (corresponding to the resistance gene of the *Agrobacterium* strain) to an optical density at 660 nm (OD-660) of approximately 0.8, centrifuged and resuspended in the LS-substrate to an OD-660 of 0.5.

The segments are left in the suspension of *Agrobacterium* for 30 minutes and then the excess of bacteria are removed by blotting the segments on sterile filter paper.

Co-cultivation

The shoot segments are co-cultured with bacteria for 48 hours directly on LS-substrate containing agar (8.0 g/l), 2,4-dichlorophenoxyacetic acid (2.0 mg/l) and trans-zeatin (0.5 mg/l). The substrate and also the explants are covered with sterile filter papers, and the petri dishes are placed at 25°C and 16 h light/ 8 dark.

"Washing" procedure

After the 48 h on the co-cultivation substrate the segments are transferred to containers containing liquid LS-substrate containing 800 mg/l carbenicillin. The containers are gently shaken and by this procedure the major part of the *Agrobacterium* is either washed off the segments and/or killed.

Selection

After the washing procedure the segments are transferred to plates containing the LS-substrate, agar (8 g/l), trans-zeatin (1-5 mg/l), gibberellic acid (0.1 mg/l), carbenicillin (800 mg/l), and kanamycin sulphate (50-100 mg/l) or phosphinotricin (1-5 mg/l) or mannose (5 g/l) depending on the vector construction used. The segments are sub-cultured to fresh substrate each 3-4 weeks.

In 3 to 4 weeks, shoots develop from the segments and the formation of new shoots continued for 3-4 months.

Rooting of regenerated shoots

The regenerated shoots are transferred to rooting substrate composed of LS-substrate, agar (8 g/l) and carbenicillin (800 mg/l).

The transgenic genotype of the regenerated shoot is verified by testing the rooting ability on the above mentioned substrates containing kanamycin sulphate (200 mg/l), by performing NPTII assays (Radke, S. E. et al, Theor. Appl. Genet. (1988), 75: 685-694) or by performing PCR analysis according to Wang *et al* (1993, NAR 21 pp 4153-4154). Plants which are not positive in any of these assays are discarded or used as controls. Alternatively, the transgenic plants could be verified by performing a GUS assay on the co-introduced β-glucuronidase gene according to Hodal, L. *et al.* (Pl. Sci. (1992), 87: 115-122).

Transfer to soil

The newly rooted plants (height approx. 2-3 cms) are transplanted from rooting substrate to soil and placed in a growth chamber (21°C, 16 hour light 200-400uE/m^2/sec). When the plants are well established they are transferred to the greenhouse, where they are grown until tubers had developed and the upper part of the plants are senescing.

Harvesting

The potatoes are harvested after about 3 months and then analysed.

BRANCHING ENZYME ANALYSIS

The class A and class B SBE expression in the transgenic potato lines is measured using the SBE assays described by Blennow and Johansson (Phytochemistry (1991) 30:437-444) and by standard Western procedures using antibodies directed against potato SBE.

STARCH ANALYSIS

Starch is isolated from potato tubers and analysed for the amylose:amylopectin ratio (Hovenkamp-Hermelink et al. (1988) Potato Research 31:241-246). In addition, the chain length distribution of amylopectin is determined by analysis of isoamylase digested starch on a Dionex HPAEC.

The number of reducing ends in isoamylase digested starch is determined by the method described by N. Nelson (1944) J. Biol.Chem. 153:375-380.

The results reveal that there is a reduction in the level of synthesis of SBE and/or the level of activity of SBE and/or the composition of starch SBE in the transgenic plants.

CONSTRUCTION OF SBE PROMOTER CONSTRUCT

An SBE promoter fragment is amplified from λ-SBE 3.4 using primers:

5' CCA TCG ATA CTT TAA GTG ATT TGA TGG C 3'

(SEQ. ID. No. 36)

and

5' CGG GAT CCT GTT CTG ATT CTT GAT TTC C 3'.

(SEQ. ID. No. 37)

The PCR product is digested with *Cla*I and *BamH*I. The resultant 1.2 kb fragment is then inserted in pVictor5a (see Figure 11) linearised with *Cla*I and *Bgl*II yielding pBEP2 (see Figure 10).

STARCH BRANCHING ENZYME MEASUREMENTS OF POTATO TUBERS

Potatoes from potato plants transformed with pBEA8, pBEA9, pSS17 or pSS18 are cut in small pieces and homogenised in extraction buffer (50 mM Tris-HCl pH 7.5, Sodium-dithionite (0.1 g/l), and 2 mM DTT) using a Ultra-Turax homogenizer; 1 g of Dowex xl. is added pr. 10 g of tuber. The crude homogenate is filtered through a miracloth filter and centrifuged at 4°C for 10 minutes at 24.700 g. The supernatant is used for starch branching enzyme assays.

The starch branching enzyme assays are carried out at 25°C in a volume of 400 μl composed of 0.1 M Na citrate buffer pH 7.0, 0.75 mg/ml amylose, 5 mg/ml bovine serum albumin and the potato extract. At 0, 15, 30 and 60 minutes aliqouts of 50 μl are

removed from the reaction into 20 µl 3 N HCl. 1 ml of iodine solution is added and the decrease in absorbance at 620 nm is measured with an ELISA spectrophotometer.

The starch branching enzyme (SBE)-levels are measured in tuber extracts from 34 transgenic Dianella potato plants transformed with plasmid pBEA8, pSS17 and pSS18.

The transformed transgenic lines produce tubers which have SBE levels that are 10% to 15% of the appropriate class A or class B SBE levels found in non transformed Dianella plants.

In a further experiment, plasmids pSS17 and pBEA8 are cotransfected into potato plants, as described above. In the cotransfectants, when analysed as set forth above, simultaneous reduction of class A and class B SBE levels are observed.

SUMMATION

The above-mentioned examples relate to the isolation, sequencing and utilisation of antisense intron constructs derived from a gene for potato class A and class B SBE. These SBE intron antisense constructs can be introduced into plants, such as potato plants. After introduction, a reduction in the level of synthesis of SBE and/or the level of activity of SBE and/or the composition of starch in plants can be achieved.

Without wishing to be bound by theory it is believed that the expressed anti-sense nucleotide sequence of the present invention binds to sense introns on pre-mRNA and thereby prevents pre-mRNA splicing and/or subsequent translation of mRNA. This binding therefore is believed to reduce the level of plant enzyme activity (in particular class A and class B SBE activity), which in turn for SBE activity is believed to influence the amylose:amylopectin ratio and thus the branching pattern of amylopectin.

Thus, the present invention provides a method wherein it is possible to manipulate the starch composition in plants, or tissues or cells thereof, such as potato tubers, by reducing the level of SBE activity by using an antisense-RNA technique using antisense intron sequences.

The simultaneous reduction or elimination of class A and class B SBE sequences from the doubly transformed potato plants, moreover, offers the possibility to transform such plants with different SBE genes at will, thus allowing the manipulation of branching in starch according to the desired result.

Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the present invention.

The following pages present a number of sequence listings which have been consecutively numbered from SEQ.I.D. No. 1 - SEQ.I.D. No. 38. In brief, SEQ.I.D. No. 1 - SEQ.I.D. No. 13 represent sense intron sequences (genomic DNA); SEQ.I.D. No. 14 represents the SBE promoter sequence (genomic sequence); SEQ.I.D. No. 15 - SEQ.I.D. No. 27 represent antisense intron sequences; and SEQ. I.D. No. 28 represents is the sequence complementary to the SBE promoter sequence - i.e. the SBE promoter sequence in antisense orientation. The full genomic nucleotide sequence for class B SBE including the promoter, exons and introns is shown as SEQ. I.D. No. 29 and is explained by way of Figures 4 and 12 which highlight particular gene features. SEQ. ID. No. 30 to 37 show primers used in the methods set forth above. SEQ. ID. No. 38 shows the sequence of intron 1 of class A SBE.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: DANISCO A/S
(B) STREET: LANGEBROGADE 1
(C) CITY: COPENHAGEN K
(E) COUNTRY: DENMARK
(F) POSTAL CODE (ZIP): DK-1001

(ii) TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION

(iii) NUMBER OF SEQUENCES: 38

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1165 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
GTAATTTTTA CTAATTTCAT GTTAATTTCA ATTATTTTTA GCCTTTGCAT TTCATTTTCC      60
AATATATCTG GATCATCTCC TTAGTTTTTT ATTTTATTTT TTATAATATC AAATATGGAA     120
GAAAAATGAC ACTTGTAGAG CCATATGTAA GTATCATGTG ACAAATTTGC AAGGTGGTTG     180
AGTGTATAAA ATTCAAAAAT TGAGAGATGG AGGGGGGGTG GGGGAAGACA ATATTTAGAA     240
AGAGTGTTCT AGGAGGTTAT GGAGGACACG GATGAGGGGT AGAAGGTTAG TTAGGTATTT     300
GAGTGTTGTC TGGCTTATCC TTTCATACTA GTAGTCGTGG AATTATTTGG GTAGTTTCTT     360
GTTTTGTTAT TTGATCTTTG TTATTCTATT TTCTGTTTCT TGTACTTCGA TTATTGTATT     420
ATATATCTTG TCGTAGTTAT TGTTCCTCGG TAAGAATGCT CTAGCATGCT TCCTTTAGTG     480
```

TTTTATCATG CCTTCTTTAT ATTCGCGTTG CTTTGAAATG CTTTTACTTT AGCCGAGGGT 540

CTATTAGAAA CAATCTCTCT ATCTCGTAAG GTAGGGGTAA AGTCCTCACC ACACTCCACT 600

TGTGGGATTA CATTGTGTTT GTTGTTGTAA ATCAATTATG TATACATAAT AAGTGGATTT 660

TTTACAACAC AAATACATGG TCAAGGGCAA AGTTCTGAAC ACATAAAGGG TTCATTATAT 720

GTCCAGGGAT ATGATAAAAA TTGTTTCTTT GTGAAAGTTA TATAAGATTT GTTATGGCTT 780

TTGCTGGAAA CATAATAAGT TATAATGCTG AGATAGCTAC TGAAGTTTGT TTTTTCTAGC 840

CTTTTAAATG TACCAATAAT AGATTCCGTA TCGAACGAGT ATGTTTTGAT TACCTGGTCA 900

TGATGTTTCT ATTTTTTACA TTTTTTTGGT GTTGAACTGC AATTGAAAAT GTTGTATCCT 960

ATGAGACGGA TAGTTGAGAA TGTGTTCTTT GTATGGACCT TGAGAAGCTC AAACGCTACT 1020

CCAATAATTT CTATGAATTC AAATTCAGTT TATGGCTACC AGTCAGTCCA GAAATTAGGA 1080

TATGCTGCAT ATACTTGTTC AATTATACTG TAAAATTTCT TAAGTTCTCA AGATATCCAT 1140

GTAACCTCGA GAATTTCTTT GACAG 1165

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 317 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GTATGTTTGA TAATTTATAT GGTTGCATGG ATAGTATATA AATAGTTGGA AAACTTCTGG 60

ACTGGTGCTC ATGGCATATT TGATCTGTGC ACCGTGTGGA GATGTCAAAC ATGTGTTACT 120

TCGTTCCGCC AATTTATAAT ACCTTAACTT GGGAAAGACA GCTCTTTACT CCTGTGGGCA 180

TTTGTTATTT GAATTACAAT CTTTATGAGC ATGGTGTTTT CACATTATCA ACTTCTTTCA 240

TGTGGTATAT AACAGTTTTT AGCTCCGTTA ATACCTTTCT TCTTTTTGAT ATAAACTAAC 300

TGTGGTGCAT TGCTTGC 317

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
GTAACAGCCA AAAGTTGTGC TTTAGGCAGT TTGACCTTAT TTTGGAAGAT GAATTGTTTA    60
TACCTACTTT GACTTTGCTA GAGAATTTTG CATACCGGGG AGTAAGTAGT GGCTCCATTT   120
AGGTGGCACC TGGCCATTTT TTTGATCTTT TAAAAAGCTG TTTGATTGGG TCTTCAAAAA   180
AGTAGACAAG GTTTTTGGAG AAGTGACACA CCCCCGGAGT GTCAGTGGCA AAGCAAAGAT   240
TTTCACTAAG GAGATTCAAA ATATAAAAAA AGTATAGACA TAAAGAAGCT GAGGGGATTC   300
AACATGTACT ATACAAGCAT CAAATATAGT CTTAAAGCAA TTTTGTAGAA ATAAAGAAAG   360
TCTTCCTTCT GTTGCTTCAC AATTTCCTTC TATTATCATG AGTTACTCTT TCTGTTCGAA   420
ATAGCTTCCT TAATATTAAA TTCATGATAC TTTTGTTGAG ATTTAGCAGT TTTTTCTTGT   480
GTAAACTGCT CTCTTTTTTT GCAG                                           504
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 146 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
GTAGGTCCTC GTCTACTACA AAATAGTAGT TTCCATCATC ATAACAGATT TTCCTATTAA    60
```

AGCATGATGT TGCAGCATCA TTGGCTTTCT TACATGTTCT AATTGCTATT AAGGTTATGC 120

TTCTAATTAA CTCATCCACA ATGCAG 146

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GTTTTGTTAT TCATACCTTG AAGCTGAATT TTGAACACCA TCATCACAGG CATTTCGATT 60

CATGTTCTTA CTAGTCTTGT TATGTAAGAC ATTTTGAAAT GCAAAAGTTA AAATAATTGT 120

GTCTTTACTA ATTTGGACTT GATCCCATAC TCTTTCCCTT AACAAAATGA GTCAATTCTA 180

TAAGTGCTTG AGAACTTACT ACTTCAGCAA TTAAACAG 218

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 198 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GTATTTTAAA TTTATTTCTA CAACTAAATA ATTCTCAGAA CAATTGTTAG ATAGAATCCA 60

AATATATACG TCCTGAAAGT ATAAAAGTAC TTATTTTCGC CATGGGCCTT CAGAATATTG 120

GTAGCCGCTG AATATCATGA TAAGTTATTT ATCCAGTGAC ATTTTTATGT TCACTCCTAT 180

TATGTCTGCT GGATACAG 198

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
GTTTGTCTGT TTCTATTGCA TTTTAAGGTT CATATAGGTT AGCCACGGAA AATCTCACTC     60
TTTGTGAGGT AACCAGGGTT CTGATGGATT ATTCAATTTT CTCGTTTATC ATTTGTTTAT    120
TCTTTTCATG CATTGTGTTT CTTTTTCAAT ATCCCTCTTA TTTGGAGGTA ATTTTTCTCA    180
TCTATTCACT TTTAGCTTCT AACCACAG                                       208
```

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
GTATGTCTTA CATCTTTAGA TATTTTGTGA TAATTACAAT TAGTTTGGCT TACTTGAACA     60
AGATTCATTC CTCAAAATGA CCTGAACTGT TGAACATCAA AGGGGTTGAA ACATAGAGGA    120
AAACAACATG ATGAATGTTT CCATTGTCTA GGGATTTCTA TTATGTTGCT GAGAACAAAT    180
GTCATCTTAA AAAAAACATT GTTTACTTTT TTGTAGTATA GAAGATTACT GTATAGAGTT    240
TGCAAGTGTG TCTGTTTTGG AGTAATTGTG AAATGTTTGA TGAACTTGTA CAG           293
```

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
GTTCAAGTAT TTTGAATCGC AGCTTGTTAA ATAATCTAGT AATTTTTAGA TTGCTTACTT    60
GGAAGTCTAC TTGGTTCTGG GGATGATAGC TCATTTCATC TTGTTCTACT TATTTTCCAA   120
CCGAATTTCT GATTTTTGTT TCGAGATCCA AGTATTAGAT TCATTTACAC TTATTACCGC   180
CTCATTTCTA CCACTAAGGC CTTGATGAGC AGCTTAAGTT GATTCTTTGA AGCTATAGTT   240
TCAGGCTACC AATCCACAGC CTGCTATATT TGTTGGATAC TTACCTTTTC TTTACAATGA   300
AGTGATACTA ATTGAAATGG TCTAAATCTG ATATCTATAT TTCTCCGTCT TTCCTCCCCC   360
TCATGATGAA ATGCAG                                                   376
```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 172 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
GTAAAATCAT CTAAAGTTGA AAGTGTTGGG TTTATGAAGT GCTTTAATTC TATCCAAGGA    60
CAAGTAGAAA CCTTTTTACC TTCCATTTCT TGATGATGGA TTTCATATTA TTTAATCCAA   120
TAGCTGGTCA AATTCGGTAA TAGCTGTACT GATTAGTTAC TTCACTTTGC AG           172
```

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 145 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```
GTATATATGT TTTACTTATC CATGAAATTA TTGCTCTGCT TGTTTTTAAT GTACTGAACA     60

AGTTTTATGG AGAAGTAACT GAAACAAATC ATTTTCACAT TGTCTAATTT AACTCTTTTT    120

TCTGATCCTC GCATGACGAA AACAG                                          145
```

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 242 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```
GTAAGGATTT GCTTGAATAA CTTTTGATAA TAAGATAACA GATGTAGGGT ACAGTTCTCT     60

CACCAAAAAG AACTGTAATT GTCTCATCCA TCTTTAGTTG TATAAGATAT CCGACTGTCT    120

GAGTTCGGAA GTGTTTGAGC CTCCTGCCCT CCCCCTGCGT TGTTTAGCTA ATTCAAAAAG    180

GAGAAAACTG TTTATTGATG ATCTTTGTCT TCATGCTGAC ATACAATCTG TTCTCATGAC    240

AG                                                                   242
```

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 797 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```
GTACAGTTCT TGCCGTGTGA CCTCCCTTTT TATTGTGGTT TTGTTCATAG TTATTTGAAT      60
GCGATAGAAG TTAACTATTG ATTACCGCCA CAATCGCCAG TTAAGTCCTC TGAACTACTA     120
ATTTGAAAGG TAGGAATAGC CGTAATAAGG TCTACTTTTG GCATCTTACT GTTACAAAAC     180
AAAAGGATGC CAAAAAAATT CTTCTCTATC CTCTTTTTCC CTAAACCAGT GCATGTAGCT     240
TGCACCTGCA TAAACTTAGG TAAATGATCA AAAATGAAGT TGATGGGAAC TTAAAACCGC     300
CCTGAAGTAA AGCTAGGAAT AGTCATATAA TGTCCACCTT TGGTGTCTGC GCTAACATCA     360
ACAACAACAT ACCTCGTGTA GTCCCACAAA GTGGTTTCAG GGGGAGGGTA GAGTGTATGC     420
AAAACTTACT CCTATCTCAG AGGTAGAGAG GATTTTTTCA ATAGACCCTT GGCTCAAGAA     480
AAAAAGTCCA AAAAGAAGTA ACAGAAGTGA AAGCAACATG TGTAGCTAAA GCGACCCAAC     540
TTGTTTGGGA CTGAAGTAGT TGTTGTTGTT GAAACAGTGC ATGTAGATGA ACACATGTCA     600
GAAAATGGAC AACACAGTTA TTTTGTGCAA GTCAAAAAAA TGTACTACTA TTTCTTTGTG     660
CAGCTTTATG TATAGAAAAG TTAAATAACT AATGAATTTT GCTAGCAGAA AAATAGCTTG     720
GAGAGAAATT TTTTATATTG AACTAAGCTA ACTATATTCA TCTTTCTTTT TGCTTCTTCT     780
TCTCCTTGTT TGTGAAG                                                    797
```

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2169 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```
ATCATGGCCA ATTACTGGTT CAAATGCATT ACTTCCTTTC AGATTCTTTC GAGTTCTCAT     60
GACCGGTCCT ACTACAGACG ATACTAACCC GTGGAACTGT TGCATCTGCT TCTTAGAACT    120
CTATGGCTAT TTTCGTTAGC TTGGCGTCGG TTTGAACATA GTTTTTGTTT TCAAACTCTT    180
CATTTACAGT CAAAATGTTG TATGGTTTTT GTTTTCCTCA ATGATGTTTA CAGTGTTGTG    240
TTGTCATCTG TACTTTTGCC TATTACTTGT TTTGAGTTAC ATGTTAAAAA AGTGTTTATT    300
TTGCCATATT TTGTTCTCTT ATTATTATTA TCATACATAC ATTATTACAA GGAAAAGACA    360
AGTACACAGA TCTTAACGTT TATGTTCAAT CAACTTTTGG AGGCATTGAC AGGTACCACA    420
AATTTTGAGT TTATGATTAA GTTCAATCTT AGAATATGAA TTTAACATCT ATTATAGATG    480
CATAAAAATA GCTAATGATA GAACATTGAC ATTTGGCAGA GCTTAGGGTA TGGTATATCC    540
AACGTTAATT TAGTAATTTT TGTTACGTAC GTATATGAAA TATTGAATTA ATCACATGAA    600
CGGTGGATAT TATATTATGA GTTGGCATCA GCAAAATCAT TGGTGTAGTT GACTGTAGTT    660
GCAGATTTAA TAATAAAATG GTAATTAACG GTCGATATTA AAATAACTCT CATTTCAAGT    720
GGGATTAGAA CTAGTTATTA AAAAAATGTA TACTTTAAGT GATTTGATGG CATATAATTT    780
AAAGTTTTTC ATTTCATGCT AAAATTGTTA ATTATTGTAA TGTAGACTGC GACTGGAATT    840
ATTATAGTGT AAATTTATGC ATTCAGTGTA AAATTAAAGT ATTGAACTTG TCTGTTTTAG    900
AAAATACTTT ATACTTTAAT ATAGGATTTT GTCATGCGAA TTTAAATTAA TCGATATTGA    960
ACACGGAATA CCAAAATTAA AAAGGATACA CATGGCCTTC ATATGAACCG TGAACCTTTG   1020
ATAACGTGGA AGTTCAAAGA AGGTAAAGTT TAAGAATAAA CTGACAAATT AATTTCTTTT   1080
ATTTGGCCCA CTACTAAATT TGCTTTACTT TCTAACATGT CAAGTTGTGC CCTCTTAGTT   1140
GAATGATATT CATTTTTCAT CCCATAAGTT CAATTTGATT GTCATACCAC CCATGATGTT   1200
CTGAAAAATG CTTGGCCATT CACAAAGTTT ATCTTAGTTC CTATGAACTT TATAAGAAGC   1260
TTTAATTTGA CATGTTATTT ATATTAGATG ATATAATCCA TGACCCAATA GACAAGTGTA   1320
TTAATATTGT AACTTTGTAA TTGAGTGTGT CTACATCTTA TTCAATCATT TAAGGTCATT   1380
AAAATAAATT ATTTTTTGAC ATTCTAAAAC TTTAAGCAGA ATAAATAGTT TATCAATTAT   1440
TAAAAACAAA AAACGACTTA TTTATAAATC AACAAACAAT TTTAGATTGC TCCAACATAT   1500
```

TTTTCCAAAT TAAATGCAGA AAATGCATAA TTTTATACTT GATCTTTATA GCTTATTTTT 1560

TTTAGCCTAA CCAACGAATA TTTGTAAACT CACAACTTGA TTAAAAGGGA TTTACAACAA 1620

GATATATATA AGTAGTGACA AATCTTGATT TTAAATATTT TAATTTGGAG GTCAAAATTT 1680

TACCATAATC ATTTGTATTT ATAATTAAAT TTTAAATATC TTATTTATAC ATATCTAGTA 1740

AACTTTTAAA TATACGTATA TACAAAATAT AAAATTATTG GCGTTCATAT TAGGTCAATA 1800

AATCCTTAAC TATATCTGCC TTACCACTAG GAGAAAGTAA AAAACTCTTT ACCAAAAATA 1860

CATGTATTAT GTATACAAAA AGTCGATTAG ATTACCTAAA TAGAAATTGT ATAACGAGTA 1920

AGTAAGTAGA AATATAAAAA AACTACAATA CTAAAAAAAA TATGTTTTAC TTCAATTTCG 1980

AAACTAATGG GGTCTGAGTG AAATATTCAG AAAGGGGAGG ACTAACAAAA GGGTCATAAT 2040

GTTTTTTTAT AAAAAGCCAC TAAAATGAGG AAATCAAGAA TCAGAACATA CAAGAAGGCA 2100

GCAGCTGAAG CAAAGTACCA TAATTTAATC AATGGAAATT AATTTCAAAG TTTTATCAAA 2160

ACCCATTCG 2169

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1165 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CTGTCAAAGA AATTCTCGAG GTTACATGGA TATCTTGAGA ACTTAAGAAA TTTTACAGTA 60

TAATTGAACA AGTATATGCA GCATATCCTA ATTTCTGGAC TGACTGGTAG CCATAAACTG 120

AATTTGAATT CATAGAAATT ATTGGAGTAG CGTTTGAGCT TCTCAAGGTC CATACAAAGA 180

ACACATTCTC AACTATCCGT CTCATAGGAT ACAACATTTT CAATTGCAGT TCAACACCAA 240

AAAAATGTAA AAAATAGAAA CATCATGACC AGGTAATCAA AACATACTCG TTCGATACGG 300

AATCTATTAT TGGTACATTT AAAAGGCTAG AAAAAACAAA CTTCAGTAGC TATCTCAGCA 360

TTATAACTTA TTATGTTTCC AGCAAAAGCC ATAACAAATC TTATATAACT TTCACAAAGA 420

AACAATTTTT ATCATATCCC TGGACATATA ATGAACCCTT TATGTGTTCA GAACTTTGCC 480

CTTGACCATG TATTTGTGTT GTAAAAAATC CACTTATTAT GTATACATAA TTGATTTACA 540

ACAACAAACA CAATGTAATC CCACAAGTGG AGTGTGGTGA GGACTTTACC CCTACCTTAC 600

GAGATAGAGA GATTGTTTCT AATAGACCCT CGGCTAAAGT AAAAGCATTT CAAAGCAACG 660

CGAATATAAA GAAGGCATGA TAAAACACTA AAGGAAGCAT GCTAGAGCAT TCTTACCGAG 720

GAACAATAAC TACGACAAGA TATATAATAC AATAATCGAA GTACAAGAAA CAGAAAATAG 780

AATAACAAAG ATCAAATAAC AAAACAAGAA ACTACCCAAA TAATTCCACG ACTACTAGTA 840

TGAAAGGATA AGCCAGACAA CACTCAAATA CCTAACTAAC CTTCTACCCC TCATCCGTGT 900

CCTCCATAAC CTCCTAGAAC ACTCTTTCTA AATATTGTCT TCCCCCACCC CCCCTCCATC 960

TCTCAATTTT TGAATTTTAT ACACTCAACC ACCTTGCAAA TTTGTCACAT GATACTTACA 1020

TATGGCTCTA CAAGTGTCAT TTTTCTTCCA TATTTGATAT TATAAAAAAT AAAATAAAAA 1080

ACTAAGGAGA TGATCCAGAT ATATTGGAAA ATGAAATGCA AAGGCTAAAA ATAATTGAAA 1140

TTAACATGAA ATTAGTAAAA ATTAC 1165

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 317 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GCAAGCAATG CACCACAGTT AGTTTATATC AAAAAGAAGA AAGGTATTAA CGGAGCTAAA 60

AACTGTTATA TACCACATGA AAGAAGTTGA TAATGTGAAA ACACCATGCT CATAAAGATT 120

GTAATTCAAA TAACAAATGC CCACAGGAGT AAAGAGCTGT CTTTCCCAAG TTAAGGTATT 180

ATAAATTGGC GGAACGAAGT AACACATGTT TGACATCTCC ACACGGTGCA CAGATCAAAT 240

ATGCCATGAG CACCAGTCCA GAAGTTTTCC AACTATTTAT ATACTATCCA TGCAACCATA 300

TAAATTATCA AACATAC 317

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CTGCAAAAAA AGAGAGCAGT TTACACAAGA AAAAACTGCT AAATCTCAAC AAAAGTATCA 60

TGAATTTAAT ATTAAGGAAG CTATTTCGAA CAGAAAGAGT AACTCATGAT AATAGAAGGA 120

AATTGTGAAG CAACAGAAGG AAGACTTTCT TTATTTCTAC AAAATTGCTT TAAGACTATA 180

TTTGATGCTT GTATAGTACA TGTTGAATCC CCTCAGCTTC TTTATGTCTA TACTTTTTTT 240

ATATTTTGAA TCTCCTTAGT GAAAATCTTT GCTTTGCCAC TGACACTCCG GGGGTGTGTC 300

ACTTCTCCAA AAACCTTGTC TACTTTTTTG AAGACCCAAT CAAACAGCTT TTTAAAAGAT 360

CAAAAAAATG GCCAGGTGCC ACCTAAATGG AGCCACTACT TACTCCCCGG TATGCAAAAT 420

TCTCTAGCAA AGTCAAAGTA GGTATAAACA ATTCATCTTC CAAAATAAGG TCAAACTGCC 480

TAAAGCACAA CTTTTGGCTG TTAC 504

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 146 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CTGCATTGTG GATGAGTTAA TTAGAAGCAT AACCTTAATA GCAATTAGAA CATGTAAGAA 60

AGCCAATGAT GCTGCAACAT CATGCTTTAA TAGGAAAATC TGTTATGATG ATGGAAACTA 120

CTATTTTGTA GTAGACGAGG ACCTAC 146

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CTGTTTAATT GCTGAAGTAG TAAGTTCTCA AGCACTTATA GAATTGACTC ATTTTGTTAA 60

GGGAAAGAGT ATGGGATCAA GTCCAAATTA GTAAAGACAC AATTATTTTA ACTTTTGCAT 120

TTCAAAATGT CTTACATAAC AAGACTAGTA AGAACATGAA TCGAAATGCC TGTGATGATG 180

GTGTTCAAAA TTCAGCTTCA AGGTATGAAT AACAAAAC 218

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 198 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CTGTATCCAG CAGACATAAT AGGAGTGAAC ATAAAAATGT CACTGGATAA ATAACTTATC 60

ATGATATTCA GCGGCTACCA ATATTCTGAA GGCCCATGGC GAAAATAAGT ACTTTTATAC 120

TTTCAGGACG TATATATTTG GATTCTATCT AACAATTGTT CTGAGAATTA TTTAGTTGTA 180

GAAATAAATT TAAAATAC 198

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CTGTGGTTAG AAGCTAAAAG TGAATAGATG AGAAAAATTA CCTCCAAATA AGAGGGATAT 60

TGAAAAAGAA ACACAATGCA TGAAAAGAAT AAACAAATGA TAAACGAGAA AATTGAATAA 120

TCCATCAGAA CCCTGGTTAC CTCACAAAGA GTGAGATTTT CCGTGGCTAA CCTATATGAA 180

CCTTAAAATG CAATAGAAAC AGACAAAC 208

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CTGTACAAGT TCATCAAACA TTTCACAATT ACTCCAAAAC AGACACACTT GCAAACTCTA 60

TACAGTAATC TTCTATACTA CAAAAAAGTA AACAATGTTT TTTTTAAGAT GACATTTGTT 120

CTCAGCAACA TAATAGAAAT CCCTAGACAA TGGAAACATT CATCATGTTG TTTTCCTCTA 180

TGTTTCAACC CCTTTGATGT TCAACAGTTC AGGTCATTTT GAGGAATGAA TCTTGTTCAA 240

GTAAGCCAAA CTAATTGTAA TTATCACAAA ATATCTAAAG ATGTAAGACA TAC 293

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CTGCATTTCA TCATGAGGGG GAGGAAAGAC GGAGAAATAT AGATATCAGA TTTAGACCAT 60

TTCAATTAGT ATCACTTCAT TGTAAAGAAA AGGTAAGTAT CCAACAAATA TAGCAGGCTG 120

TGGATTGGTA GCCTGAAACT ATAGCTTCAA AGAATCAACT TAAGCTGCTC ATCAAGGCCT 180

TAGTGGTAGA AATGAGGCGG TAATAAGTGT AAATGAATCT AATACTTGGA TCTCGAAACA 240

AAAATCAGAA ATTCGGTTGG AAAATAAGTA GAACAAGATG AAATGAGCTA TCATCCCCAG 300

AACCAAGTAG ACTTCCAAGT AAGCAATCTA AAAATTACTA GATTATTTAA CAAGCTGCGA 360

TTCAAAATAC TTGAAC 376

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 172 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CTGCAAAGTG AAGTAACTAA TCAGTACAGC TATTACCGAA TTTGACCAGC TATTGGATTA 60

AATAATATGA AATCCATCAT CAAGAAATGG AAGGTAAAAA GGTTTCTACT TGTCCTTGGA 120

TAGAATTAAA GCACTTCATA AACCCAACAC TTTCAACTTT AGATGATTTT AC 172

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 145 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CTGTTTTCGT CATGCGAGGA TCAGAAAAAA GAGTTAAATT AGACAATGTG AAAATGATTT 60

GTTTCAGTTA CTTCTCCATA AAACTTGTTC AGTACATTAA AAACAAGCAG AGCAATAATT 120

TCATGGATAA GTAAAACATA TATAC 145

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 242 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CTGTCATGAG AACAGATTGT ATGTCAGCAT GAAGACAAAG ATCATCAATA AACAGTTTTC 60

TCCTTTTTGA ATTAGCTAAA CAACGCAGGG GGAGGGCAGG AGGCTCAAAC ACTTCCGAAC 120

TCAGACAGTC GGATATCTTA TACAACTAAA GATGGATGAG ACAATTACAG TTCTTTTTGG 180

TGAGAGAACT GTACCCTACA TCTGTTATCT TATTATCAAA AGTTATTCAA GCAAATCCTT 240

AC 242

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 797 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

```
CTTCACAAAC AAGGAGAAGA AGAAGCAAAA AGAAAGATGA ATATAGTTAG CTTAGTTCAA      60
TATAAAAAAT TTCTCTCCAA GCTATTTTTC TGCTAGCAAA ATTCATTAGT TATTTAACTT     120
TTCTATACAT AAAGCTGCAC AAAGAAATAG TAGTACATTT TTTTGACTTG CACAAAATAA     180
CTGTGTTGTC CATTTTCTGA CATGTGTTCA TCTACATGCA CTGTTTCAAC AACAACAACT     240
ACTTCAGTCC CAAACAAGTT GGGTCGCTTT AGCTACACAT GTTGCTTTCA CTTCTGTTAC     300
TTCTTTTTGG ACTTTTTTTC TTGAGCCAAG GGTCTATTGA AAAAATCCTC TCTACCTCTG     360
AGATAGGAGT AAGTTTTGCA TACACTCTAC CCTCCCCCTG AAACCACTTT GTGGGACTAC     420
ACGAGGTATG TTGTTGTTGA TGTTAGCGCA GACACCAAAG GTGGACATTA TATGACTATT     480
CCTAGCTTTA CTTCAGGGCG GTTTTAAGTT CCCATCAACT TCATTTTTGA TCATTTACCT     540
AAGTTTATGC AGGTGCAAGC TACATGCACT GGTTTAGGGA AAAAGAGGAT AGAGAAGAAT     600
TTTTTTGGCA TCCTTTTGTT TTGTAACAGT AAGATGCCAA AAGTAGACCT TATTACGGCT     660
ATTCCTACCT TTCAAATTAG TAGTTCAGAG GACTTAACTG GCGATTGTGG CGGTAATCAA     720
TAGTTAACTT CTATCGCATT CAAATAACTA TGAACAAAAC CACAATAAAA AGGGAGGTCA     780
CACGGCAAGA ACTGTAC                                                    797
```

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2169 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

```
CGAATGGGTT TTGATAAAAC TTTGAAATTA ATTTCCATTG ATTAAATTAT GGTACTTTGC     60
TTCAGCTGCT GCCTTCTTGT ATGTTCTGAT TCTTGATTTC CTCATTTTAG TGGCTTTTTA    120
TAAAAAAACA TTATGACCCT TTTGTTAGTC CTCCCCTTTC TGAATATTTC ACTCAGACCC    180
CATTAGTTTC GAAATTGAAG TAAAACATAT TTTTTTTAGT ATTGTAGTTT TTTTATATTT    240
CTACTTACTT ACTCGTTATA CAATTTCTAT TTAGGTAATC TAATCGACTT TTTGTATACA    300
TAATACATGT ATTTTTGGTA AAGAGTTTTT TACTTTCTCC TAGTGGTAAG GCAGATATAG    360
TTAAGGATTT ATTGACCTAA TATGAACGCC AATAATTTTA TATTTTGTAT ATACGTATAT    420
TTAAAAGTTT ACTAGATATG TATAAATAAG ATATTTAAAA TTTAATTATA AATACAAATG    480
ATTATGGTAA AATTTTGACC TCCAAATTAA AATATTTAAA ATCAAGATTT GTCACTACTT    540
ATATATATCT TGTTGTAAAT CCCTTTTAAT CAAGTTGTGA GTTTACAAAT ATTCGTTGGT    600
TAGGCTAAAA AAAATAAGCT ATAAAGATCA AGTATAAAAT TATGCATTTT CTGCATTTAA    660
TTTGGAAAAA TATGTTGGAG CAATCTAAAA TTGTTTGTTG ATTTATAAAT AAGTCGTTTT    720
TTGTTTTTAA TAATTGATAA ACTATTTATT CTGCTTAAAG TTTTAGAATG TCAAAAAATA    780
ATTTATTTTA ATGACCTTAA ATGATTGAAT AAGATGTAGA CACACTCAAT TACAAAGTTA    840
CAATATTAAT ACACTTGTCT ATTGGGTCAT GGATTATATC ATCTAATATA AATAACATGT    900
CAAATTAAAG CTTCTTATAA AGTTCATAGG AACTAAGATA AACTTTGTGA ATGGCCAAGC    960
ATTTTTCAGA ACATCATGGG TGGTATGACA ATCAAATTGA ACTTATGGGA TGAAAAATGA   1020
ATATCATTCA ACTAAGAGGG CACAACTTGA CATGTTAGAA AGTAAAGCAA ATTTAGTAGT   1080
GGGCCAAATA AAAGAAATTA ATTTGTCAGT TTATTCTTAA ACTTTACCTT CTTTGAACTT   1140
CCACGTTATC AAAGGTTCAC GGTTCATATG AAGGCCATGT GTATCCTTTT TAATTTTGGT   1200
ATTCCGTGTT CAATATCGAT TAATTTAAAT TCGCATGACA AAATCCTATA TTAAAGTATA   1260
AAGTATTTTC TAAAACAGAC AAGTTCAATA CTTTAATTTT ACACTGAATG CATAAATTTA   1320
CACTATAATA ATTCCAGTCG CAGTCTACAT TACAATAATT AACAATTTTA GCATGAAATG   1380
AAAAACTTTA AATTATATGC CATCAAATCA CTTAAAGTAT ACATTTTTTT AATAACTAGT   1440
TCTAATCCCA CTTGAAATGA GAGTTATTTT AATATCGACC GTTAATTACC ATTTTATTAT   1500
```

TAAATCTGCA ACTACAGTCA ACTACACCAA TGATTTTGCT GATGCCAACT CATAATATAA 1560

TATCCACCGT TCATGTGATT AATTCAATAT TTCATATACG TACGTAACAA AAATTACTAA 1620

ATTAACGTTG GATATACCAT ACCCTAAGCT CTGCCAAATG TCAATGTTCT ATCATTAGCT 1680

ATTTTTATGC ATCTATAATA GATGTTAAAT TCATATTCTA AGATTGAACT TAATCATAAA 1740

CTCAAAATTT GTGGTACCTG TCAATGCCTC CAAAAGTTGA TTGAACATAA ACGTTAAGAT 1800

CTGTGTACTT GTCTTTTCCT TGTAATAATG TATGTATGAT AATAATAATA AGAGAACAAA 1860

ATATGGCAAA ATAAACACTT TTTTAACATG TAACTCAAAA CAAGTAATAG GCAAAAGTAC 1920

AGATGACAAC ACAACACTGT AAACATCATT GAGGAAAACA AAAACCATAC AACATTTTGA 1980

CTGTAAATGA AGAGTTTGAA AACAAAAACT ATGTTCAAAC CGACGCCAAG CTAACGAAAA 2040

TAGCCATAGA GTTCTAAGAA GCAGATGCAA CAGTTCCACG GGTTAGTATC GTCTGTAGTA 2100

GGACCGGTCA TGAGAACTCG AAAGAATCTG AAAGGAAGTA ATGCATTTGA ACCAGTAATT 2160

GGCCATGAT 2169

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ATCATGGCCA ATTACTGGTT CAAATGCATT ACTTCCTTTC AGATTCTTTC GAGTTCTCAT 60

GACCGGTCCT ACTACAGACG ATACTAACCC GTGGAACTGT TGCATCTGCT TCTTAGAACT 120

CTATGGCTAT TTTCGTTAGC TTGGCGTCGG TTTGAACATA GTTTTTGTTT TCAAACTCTT 180

CATTTACAGT CAAAATGTTG TATGGTTTTT GTTTTCCTCA ATGATGTTTA CAGTGTTGTG 240

TTGTCATCTG TACTTTTGCC TATTACTTGT TTTGAGTTAC ATGTTAAAAA AGTGTTTATT 300

TTGCCATATT TTGTTCTCTT ATTATTATTA TCATACATAC ATTATTACAA GGAAAAGACA 360

AGTACACAGA TCTTAACGTT TATGTTCAAT CAACTTTTGG AGGCATTGAC AGGTACCACA 420

AATTTTGAGT TTATGATTAA GTTCAATCTT AGAATATGAA TTTAACATCT ATTATAGATG 480

CATAAAAATA GCTAATGATA GAACATTGAC ATTTGGCAGA GCTTAGGGTA TGGTATATCC 540

AACGTTAATT TAGTAATTTT TGTTACGTAC GTATATGAAA TATTGAATTA ATCACATGAA 600

CGGTGGATAT TATATTATGA GTTGGCATCA GCAAAATCAT TGGTGTAGTT GACTGTAGTT 660

GCAGATTTAA TAATAAAATG GTAATTAACG GTCGATATTA AAATAACTCT CATTTCAAGT 720

GGGATTAGAA CTAGTTATTA AAAAAATGTA TACTTTAAGT GATTTGATGG CATATAATTT 780

AAAGTTTTTC ATTTCATGCT AAAATTGTTA ATTATTGTAA TGTAGACTGC GACTGGAATT 840

ATTATAGTGT AAATTTATGC ATTCAGTGTA AAATTAAAGT ATTGAACTTG TCTGTTTTAG 900

AAAATACTTT ATACTTTAAT ATAGGATTTT GTCATGCGAA TTTAAATTAA TCGATATTGA 960

ACACGGAATA CCAAAATTAA AAAGGATACA CATGGCCTTC ATATGAACCG TGAACCTTTG 1020

ATAACGTGGA AGTTCAAAGA AGGTAAAGTT TAAGAATAAA CTGACAAATT AATTTCTTTT 1080

ATTTGGCCCA CTACTAAATT TGCTTTACTT TCTAACATGT CAAGTTGTGC CCTCTTAGTT 1140

GAATGATATT CATTTTTCAT CCCATAAGTT CAATTTGATT GTCATACCAC CCATGATGTT 1200

CTGAAAAATG CTTGGCCATT CACAAAGTTT ATCTTAGTTC CTATGAACTT TATAAGAAGC 1260

TTTAATTTGA CATGTTATTT ATATTAGATG ATATAATCCA TGACCCAATA GACAAGTGTA 1320

TTAATATTGT AACTTTGTAA TTGAGTGTGT CTACATCTTA TTCAATCATT TAAGGTCATT 1380

AAAATAAATT ATTTTTTGAC ATTCTAAAAC TTTAAGCAGA ATAAATAGTT TATCAATTAT 1440

TAAAAACAAA AAACGACTTA TTTATAAATC AACAAACAAT TTTAGATTGC TCCAACATAT 1500

TTTTCCAAAT TAAATGCAGA AAATGCATAA TTTTATACTT GATCTTTATA GCTTATTTTT 1560

TTTAGCCTAA CCAACGAATA TTTGTAAACT CACAACTTGA TTAAAAGGGA TTTACAACAA 1620

GATATATATA AGTAGTGACA AATCTTGATT TTAAATATTT TAATTTGGAG GTCAAAATTT 1680

TACCATAATC ATTTGTATTT ATAATTAAAT TTTAAATATC TTATTTATAC ATATCTAGTA 1740

AACTTTTAAA TATACGTATA TACAAAATAT AAAATTATTG GCGTTCATAT TAGGTCAATA 1800

AATCCTTAAC TATATCTGCC TTACCACTAG GAGAAAGTAA AAAACTCTTT ACCAAAAATA 1860

CATGTATTAT GTATACAAAA AGTCGATTAG ATTACCTAAA TAGAAATTGT ATAACGAGTA 1920

AGTAAGTAGA AATATAAAAA AACTACAATA CTAAAAAAAA TATGTTTTAC TTCAATTTCG 1980

AAACTAATGG GGTCTGAGTG AAATATTCAG AAAGGGGAGG ACTAACAAAA GGGTCATAAT 2040

GTTTTTTTAT AAAAAGCCAC TAAAATGAGG AAATCAAGAA TCAGAACATA CAAGAAGGCA 2100

GCAGCTGAAG CAAAGTACCA TAATTTAATC AATGGAAATT AATTTCAAAG TTTTATCAAA 2160
ACCCATTCGA GGATCTTTTC CATCTTTCTC ACCTAAAGTT TCTTCAGGGG TAATTTTTAC 2220
TAATTTCATG TTAATTTCAA TTATTTTTAG CCTTTGCATT TCATTTTCCA ATATATCTGG 2280
ATCATCTCCT TAGTTTTTTA TTTTATTTTT TATAATATCA AATATGGAAG AAAAATGACA 2340
CTTGTAGAGC CATATGTAAG TATCATGTGA CAAATTTGCA AGGTGGTTGA GTGTATAAAA 2400
TTCAAAAATT GAGAGATGGA GGGGGGGTGG GGGAAGACAA TATTTAGAAA GAGTGTTCTA 2460
GGAGGTTATG GAGGACACGG ATGAGGGGTA GAAGGTTAGT TAGGTATTTG AGTGTTGTCT 2520
GGCTTATCCT TTCATACTAG TAGTCGTGGA ATTATTTGGG TAGTTTCTTG TTTTGTTATT 2580
TGATCTTTGT TATTCTATTT TCTGTTTCTT GTACTTCGAT TATTGTATTA TATATCTTGT 2640
CGTAGTTATT GTTCCTCGGT AAGAATGCTC TAGCATGCTT CCTTTAGTGT TTTATCATGC 2700
CTTCTTTATA TTCGCGTTGC TTTGAAATGC TTTTACTTTA GCCGAGGGTC TATTAGAAAC 2760
AATCTCTCTA TCTCGTAAGG TAGGGGTAAA GTCCTCACCA CACTCCACTT GTGGGATTAC 2820
ATTGTGTTTG TTGTTGTAAA TCAATTATGT ATACATAATA AGTGGATTTT TTACAACACA 2880
AATACATGGT CAAGGGCAAA GTTCTGAACA CATAAAGGGT TCATTATATG TCCAGGGATA 2940
TGATAAAAAT TGTTTCTTTG TGAAAGTTAT ATAAGATTTG TTATGGCTTT TGCTGGAAAC 3000
ATAATAAGTT ATAATGCTGA GATAGCTACT GAAGTTTGTT TTTTCTAGCC TTTTAAATGT 3060
ACCAATAATA GATTCCGTAT CGAACGAGTA TGTTTTGATT ACCTGGTCAT GATGTTTCTA 3120
TTTTTTACAT TTTTTTGGTG TTGAACTGCA ATTGAAAATG TTGTATCCTA TGAGACGGAT 3180
AGTTGAGAAT GTGTTCTTTG TATGGACCTT GAGAAGCTCA AACGCTACTC CAATAATTTC 3240
TATGAATTCA AATTCAGTTT ATGGCTACCA GTCAGTCCAG AAATTAGGAT ATGCTGCATA 3300
TACTTGTTCA ATTATACTGT AAAATTTCTT AAGTTCTCAA GATATCCATG TAACCTCGAG 3360
AATTTCTTTG ACAGGCTTCT AGAAATAAGA TATGTTTTCC TTCTCAACAT AGTACTGGAC 3420
TGAAGTTTGG ATCTCAGGAA CGGTCTTGGG ATATTTCTTC CACCCCAAAA TCAAGAGTTA 3480
GAAAAGATGA AAGGGTATGT TTGATAATTT ATATGGTTGC ATGGATAGTA TATAAATAGT 3540
TGGAAAACTT CTGGACTGGT GCTCATGGCA TATTTGATCT GTGCACCGTG TGGAGATGTC 3600
AAACATGTGT TACTTCGTTC CGCCAATTTA TAATACCTTA ACTTGGGAAA GACAGCTCTT 3660
TACTCCTGTG GGCATTTGTT ATTTGAATTA CAATCTTTAT GAGCATGGTG TTTTCACATT 3720
ATCAACTTCT TTCATGTGGT ATATAACAGT TTTTAGCTCC GTTAATACCT TTCTTCTTTT 3780
TGATATAAAC TAACTGTGGT GCATTGCTTG CATGAAGCAC AGTTCAGCTA TTTCCGCTGT 3840

TTTGACCGAT GACGACAATT CGACAATGGC ACCCCTAGAG GAAGATGTCA AGACTGAAAA 3900

TATTGGCCTC CTAAATTTGG ATCCAACTTT GGAACCTTAT CTAGATCACT TCAGACACAG 3960

AATGAAGAGA TATGTGGATC AGAAAATGCT CATTGAAAAA TATGAGGGAC CCCTTGAGGA 4020

ATTTGCTCAA GGTAACAGCC AAAAGTTGTG CTTTAGGCAG TTTGACCTTA TTTTGGAAGA 4080

TGAATTGTTT ATACCTACTT TGACTTTGCT AGAGAATTTT GCATACCGGG GAGTAAGTAG 4140

TGGCTCCATT TAGGTGGCAC CTGGCCATTT TTTTGATCTT TTAAAAAGCT GTTTGATTGG 4200

GTCTTCAAAA AAGTAGACAA GGTTTTTGGA GAAGTGACAC ACCCCCGGAG TGTCAGTGGC 4260

AAAGCAAAGA TTTTCACTAA GGAGATTCAA AATATAAAAA AAGTATAGAC ATAAAGAAGC 4320

TGAGGGGATT CAACATGTAC TATACAAGCA TCAAATATAG TCTTAAAGCA ATTTTGTAGA 4380

AATAAAGAAA GTCTTCCTTC TGTTGCTTCA CAATTTCCTT CTATTATCAT GAGTTACTCT 4440

TTCTGTTCGA AATAGCTTCC TTAATATTAA ATTCATGATA CTTTTGTTGA GATTTAGCAG 4500

TTTTTTCTTG TGTAAACTGC TCTCTTTTTT TGCAGGTTAT TTAAAATTTG GATTCAACAG 4560

GGAAGATGGT TGCATAGTCT ATCGTGAATG GGCTCCTGCT GCTCAGTAGG TCCTCGTCTA 4620

CTACAAAATA GTAGTTTCCA TCATCATAAC AGATTTTCCT ATTAAAGCAT GATGTTGCAG 4680

CATCATTGGC TTTCTTACAT GTTCTAATTG CTATTAAGGT TATGCTTCTA ATTAACTCAT 4740

CCACAATGCA GGGAAGCAGA AGTTATTGGC GATTTCAATG GATGGAACGG TTCTAACCAC 4800

ATGATGGAGA AGGACCAGTT TGGTGTTTGG AGTATTAGAA TTCCTGATGT TGACAGTAAG 4860

CCAGTCATTC CACACAACTC CAGAGTTAAG TTTCGTTTCA AACATGGTAA TGGAGTGTGG 4920

GTAGATCGTA TCCCTGCTTG GATAAAGTAT GCCACTGCAG ACGCCACAAA GTTTGCAGCA 4980

CCATATGATG GTGTCTACTG GGACCCACCA CCTTCAGAAA GGTTTTGTTA TTCATACCTT 5040

GAAGCTGAAT TTTGAACACC ATCATCACAG GCATTTCGAT TCATGTTCTT ACTAGTCTTG 5100

TTATGTAAGA CATTTTGAAA TGCAAAAGTT AAAATAATTG TGTCTTTACT AATTTGGACT 5160

TGATCCCATA CTCTTTCCCT TAACAAAATG AGTCAATTCT ATAAGTGCTT GAGAACTTAC 5220

TACTTCAGCA ATTAAACAGG TACCACTTCA AATACCCTCG CCCTCCCAAA CCCCGAGCCC 5280

CACGAATCTA TGAAGCACAT GTCGGCATGA GCAGCTCTGA GCCACGTGTA AATTCGTATC 5340

GTGAGTTTGC AGATGATGTT TTACCTCGGA TTAAGGCAAA TAACTATAAT ACTGTCCAGT 5400

TGATGGCCAT AATGGAACAT TCTTACTATG GATCATTTGG ATATCATGTT ACAAACTTTT 5460

TTGCTGTGAG CAGTAGATAT GGAAACCCGG AGGACCTAAA GTATCTGATA GATAAAGCAC 5520

```
ATAGCTTGGG TTTACAGGTT CTGGTGGATG TAGTTCACAG TCATGCAAGC AATAATGTCA   5580
CTGATGGCCT CAATGGCTTT GATATTGGCC AAGGTTCTCA AGAATCCTAC TTTCATGCTG   5640
GAGAGCGAGG GTACCATAAG TTGTGGGATA GCAGGCTGTT CAACTATGCC AATTGGGAGG   5700
TTCTTCGTTT CCTTCTTTCC AACTTGAGGT GGTGGCTAGA AGAGTATAAC TTTGACGGAT   5760
TTCGATTTGA TGGAATAACT TCTATGCTGT ATGTTCATCA TGGAATCAAT ATGGGATTTA   5820
CAGGAAACTA TAATGAGTAT TTCAGCGAGG CTACAGATGT TGATGCTGTG GTCTATTTAA   5880
TGTTGGCCAA TAATCTGATT CACAAGATTT TCCCAGATGC AACTGTTATT GCCGAAGATG   5940
TTTCTGGTAT GCCGGGCCTT GGCCGGCCTG TTTCTGAGGG AGGAATTGGT TTTGTTTACC   6000
GCCTGGCAAT GGCAATCCCA GATAAGTGGA TAGATTATTT AAAGAATAAG AATGATGAAG   6060
ATTGGTCCAT GAAGGAAGTA ACATCGAGTT TGACAAATAG GAGATATACA GAGAAGTGTA   6120
TAGCATATGC GGAGACCCAT GATCAGGTAT TTTAAATTTA TTTCTACAAC TAAATAATTC   6180
TCAGAACAAT TGTTAGATAG AATCCAAATA TATACGTCCT GAAAGTATAA AAGTACTTAT   6240
TTTCGCCATG GGCCTTCAGA ATATTGGTAG CCGCTGAATA TCATGATAAG TTATTTATCC   6300
AGTGACATTT TTATGTTCAC TCCTATTATG TCTGCTGGAT ACAGTCTATT GTTGGTGACA   6360
AGACCATTGC ATTTCTCCTA ATGGACAAAG AGATGTATTC TGGCATGTCT TGCTTGACAG   6420
ATGCTTCTCC TGTTGTTGAT CGAGGAATTG CGCTTCACAA GGTTTGTCTG TTTCTATTGC   6480
ATTTTAAGGT TCATATAGGT TAGCCACGGA AAATCTCACT CTTTGTGAGG TAACCAGGGT   6540
TCTGATGGAT TATTCAATTT TCTCGTTTAT CATTTGTTTA TTCTTTTCAT GCATTGTGTT   6600
TCTTTTTCAA TATCCCTCTT ATTTGGAGGT AATTTTTCTC ATCTATTCAC TTTTAGCTTC   6660
TAACCACAGA TGATCCATTT TTTCACAATG GCCTTGGGAG GAGAGGGGTA CCTCAATTTC   6720
ATGGGTAACG AGGTATGTCT TACATCTTTA GATATTTTGT GATAATTACA ATTAGTTTGG   6780
CTTACTTGAA CAAGATTCAT TCCTCAAAAT GACCTGAACT GTTGAACATC AAAGGGGTTG   6840
AAACATAGAG GAAAACAACA TGATGAATGT TTCCATTGTC TAGGGATTTC TATTATGTTG   6900
CTGAGAACAA ATGTCATCTT AAAAAAAACA TTGTTTACTT TTTTGTAGTA TAGAAGATTA   6960
CTGTATAGAG TTTGCAAGTG TGTCTGTTTT GGAGTAATTG TGAAATGTTT GATGAACTTG   7020
TACAGTTTGG CCATCCTGAG TGGATTGACT TCCCTAGAGA GGGCAATAAT TGGAGTTATG   7080
ACAAATGTAG ACGCCAGTGG AACCTCGCGG ATAGCGAACA CTTGAGATAC AAGGTTCAAG   7140
TATTTTGAAT CGCAGCTTGT TAAATAATCT AGTAATTTTT AGATTGCTTA CTTGGAAGTC   7200
TACTTGGTTC TGGGGATGAT AGCTCATTTC ATCTTGTTCT ACTTATTTTC CAACCGAATT   7260
```

TCTGATTTTT GTTTCGAGAT CCAAGTATTA GATTCATTTA CACTTATTAC CGCCTCATTT 7320

CTACCACTAA GGCCTTGATG AGCAGCTTAA GTTGATTCTT TGAAGCTATA GTTTCAGGCT 7380

ACCAATCCAC AGCCTGCTAT ATTTGTTGGA TACTTACCTT TTCTTTACAA TGAAGTGATA 7440

CTAATTGAAA TGGTCTAAAT CTGATATCTA TATTTCTCCG TCTTTCCTCC CCCTCATGAT 7500

GAAATGCAGT TTATGAATGC ATTTGATAGA GCTATGAATT CGCTCGATGA AAAGTTCTCA 7560

TTCCTCGCAT CAGGAAAACA GATAGTAAGC AGCATGGATG ATGATAATAA GGTAAAATCA 7620

TCTAAAGTTG AAAGTGTTGG GTTTATGAAG TGCTTTAATT CTATCCAAGG ACAAGTAGAA 7680

ACCTTTTTAC CTTCCATTTC TTGATGATGG ATTTCATATT ATTTAATCCA ATAGCTGGTC 7740

AAATTCGGTA ATAGCTGTAC TGATTAGTTA CTTCACTTTG CAGGTTGTTG TGTTTGAACG 7800

TGGTGACCTG GTATTTGTAT TCAACTTCCA CCCAAAGAAC ACATACGAAG GGTATATATG 7860

TTTTACTTAT CCATGAAATT ATTGCTCTGC TTGTTTTTAA TGTACTGAAC AAGTTTTATG 7920

GAGAAGTAAC TGAAACAAAT CATTTTCACA TTGTCTAATT TAACTCTTTT TTCTGATCCT 7980

CGCATGACGA AAACAGGTAT AAAGTTGGAT GTGACTTGCC AGGGAAGTAC AGAGTTGCAC 8040

TGGACAGTGA TGCTTGGGAA TTTGGTGGCC ATGGAAGAGT AAGGATTTGC TTGAATAACT 8100

TTTGATAATA AGATAACAGA TGTAGGGTAC AGTTCTCTCA CCAAAAAGAA CTGTAATTGT 8160

CTCATCCATC TTTAGTTGTA TAAGATATCC GACTGTCTGA GTTCGGAAGT GTTTGAGCCT 8220

CCTGCCCTCC CCCTGCGTTG TTTAGCTAAT TCAAAAAGGA GAAAACTGTT TATTGATGAT 8280

CTTTGTCTTC ATGCTGACAT ACAATCTGTT CTCATGACAG ACTGGTCATG ATGTTGACCA 8340

TTTCACATCA CCAGAAGGAA TACCTGGAGT TCCAGAAACA AATTTCAATG GTCGTCCAAA 8400

TTCCTTCAAA GTGCTGTCTC CTGCGCGAAC ATGTGTGGTA CAGTTCTTGC CGTGTGACCT 8460

CCCTTTTTAT TGTGGTTTTG TTCATAGTTA TTTGAATGCG ATAGAAGTTA ACTATTGATT 8520

ACCGCCACAA TCGCCAGTTA AGTCCTCTGA ACTACTAATT TGAAAGGTAG GAATAGCCGT 8580

AATAAGGTCT ACTTTTGGCA TCTTACTGTT ACAAAACAAA AGGATGCCAA AAAAATTCTT 8640

CTCTATCCTC TTTTTCCCTA AACCAGTGCA TGTAGCTTGC ACCTGCATAA ACTTAGGTAA 8700

ATGATCAAAA ATGAAGTTGA TGGGAACTTA AAACCGCCCT GAAGTAAAGC TAGGAATAGT 8760

CATATAATGT CCACCTTTGG TGTCTGCGCT AACATCAACA ACAACATACC TCGTGTAGTC 8820

CCACAAAGTG GTTTCAGGGG GAGGGTAGAG TGTATGCAAA ACTTACTCCT ATCTCAGAGG 8880

TAGAGAGGAT TTTTTCAATA GACCCTTGGC TCAAGAAAAA AAGTCCAAAA AGAAGTAACA 8940

GAAGTGAAAG CAACATGTGT AGCTAAAGCG ACCCAACTTG TTTGGGACTG AAGTAGTTGT 9000

TGTTGTTGAA ACAGTGCATG TAGATGAACA CATGTCAGAA AATGGACAAC ACAGTTATTT 9060

TGTGCAAGTC AAAAAAATGT ACTACTATTT CTTTGTGCAG CTTTATGTAT AGAAAAGTTA 9120

AATAACTAAT GAATTTTGCT AGCAGAAAAA TAGCTTGGAG AGAAATTTTT TATATTGAAC 9180

TAAGCTAACT ATATTCATCT TTCTTTTTGC TTCTTCTTCT CCTTGTTTGT GAAGGCTTAT 9240

TACAGAGTTG ATGAACGCAT GTCAGAAACT GAAGATTACC AGACAGACAT TTGTAGTGAG 9300

CTACTACCAA CAGCCAATAT CGAGGAGAGT GACGAGAAAC TTAAAGATTC GTTATCTACA 9360

AATATCAGTA ACATTGACGA ACGCATGTCA GAAACTGAAG TTTACCAGAC AGACATTTCT 9420

AGTGAGCTAC TACCAACAGC CAATATTGAG GAGAGTGACG AGAAACTTAA AGATTCGTTA 9480

TCTACAAATA TCAGTAACAT TGATCAGACT GTTGTAGTTT CTGTTGAGGA GAGAGACAAG 9540

GAACTTAAAG ATTCACCGTC TGTAAGCATC ATTAGTGATG TTGTTCCAGC TGAATGGGAT 9600

GATTCAGATG CAAACGTCTG GGGTGAGGAC TAGTCAGATG ATTGATCGAC CCTTCTACCG 9660

ATTGGTGATC GCTATCCTTG CTCTCTGAGA AATAGGTGAG GCGAAACAAA AAATAATTTG 9720

CATGATAAAA AGTCTGATTT TATGATCGCT ATCCTCGCTC TCTGAGAAAG AAGCGAAACA 9780

AAGGCGACTC CTGGACTCGA ATCTATAAGA TAACAAAGGC GACTCCTGGG ACTCGAATCT 9840

ATAAGATAAC AAAGGCAATT CCAAGACTTG AATCTATAAA AAATTTAGTT AAGAATGATT 9900

AACGTCCGAT CCTAATTCGA ATCGAGGCAT CTTACCACTC CATTGATAAT TATATAAGTC 9960

AATAAGTCAT ATAAAGTATT AAAAACTAAA TTGACTTGAT CGGTCTATCA AAAATAGATA 10020

AATTGTGTTC ATATGTAACA TTTTTGTTGT CACAATTAGC TTAATTACAT CTTTCATGTG 10080

CAATAACAAA GAAATGATAG GAATTTAGAG ATTCCAATTT TTTTGTTGCC ACAATTAACT 10140

TAATTACATC TTTCATTTGC AATAACAAAG AAATGATAGG AATTTAGAGA TCCAGTGTCA 10200

ATACACAACC TAGGCCAACA TCGAAAGCAT AACTGTAAAC TCATGCATGA AGAAATCAGT 10260

CGTAAAAATG AATAAATGCG ACATAAAAAC AAATTGCATG TATCATTAAT GTGACTTAAC 10320

TACAAGTAAA AATAAATTTA ACAAATGTAA CTTAACTACA AGTAAAAATA AATTGCTTCT 10380

ATCATTAACA AACAAACAGA ATTAAAAAGA AAAAAACATA CTAAATCTTA CCGTCATTCG 10440

ATAAAAAAAA ATACCAAATT CATAATGCAA GGAAAACGAA ACGCGTCCTG ATCGGGTATC 10500

AACGATGAAA TGGACCAGTT GGATCGACTG CCTGCACAAC GTTAGGTATG CCAAAAAAAA 10560

GAACACGATC CTTTGCACCC GTTCGATGAT TATCAGTATG TTCACAAAAA AAACTTAAGT 10620

TCATCCCAGT GTACAACAGC CCCAACATCT GCCCCAAGTA ACAAAAAACA ACCAATTTAT 10680

CTTATTCTTA TCTGCCACAA AATAATCGGT TTCACACTAT TCTCTTGTTA TACAAAATTG 10740

ACAAGTAGGA AGGAGAGGAG TCATCCAAAT AAACGGTGCA CGTTCTTTGA GAAAAGTCTT 10800

ATTTTTCGTA AGATCCAATT TCAACAAACT TTTCTTCAAG TCAAAATTCC TGATAGTGTA 10860

TCTCCTCTCG ACGACCTCTT GCATTGAACG ATCTCCGCTT ATCATGAAAA GTTGCTTGGA 10920

TAACAAGTAT TGCAAGGGGG GGACAGTAGC TATTAAGTTA GTCGGCCCAA GGAAATGGAG 10980

GAGTGATAGT CTCGAATATT ATTCACCTCT TTAGCATTAC CCGGTCTGGC TTTAAGGAGT 11040

TACGTCTTTT ACGCTCGCCA ATTTCTTTTT TTAGAATGGT TGGTGTCAAA ATCGCGAGTT 11100

GTGGAAGGTT CAAGTTACTC GATTCGTGAT TTTCAAGTAT GAGTGGTGAG AGAGATTCGA 11160

TATTTTCACG AGGTGTATTC GAGGTCTAGT AGAACGAAGG GTGTCACTAA TGAAAGTTTC 11220

AAGAGTTCAT CATCATCTTC TTCTAGTAGA TTTTCGCTTT CAAATGAGTA TGAAAATTCT 11280

TCCTCTTTTC TATTGATTTT CTTCATTGTT TTCTTCATTG TTGTGGTTGT TATTGAAAAG 11340

AAAGAAAATT TATAACAGAA AAAGATGTCA AAAAAAAGGT AAAATGAAAG AGTATCATAT 11400

ACTTAAAGAG TTGCGTAGAG ATAAGTCAAA AGAAACAGAA TTATAGTAAT TTCAGCTAAG 11460

TTAGAATTC 11469

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GGAATTCCAG TCGCAGTCTA CATTAC 26

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CGGGATCCAG AGGCATTAAG ATTTCTGG 28

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CGGGATCCAA AGAAATTCTC GAGGTTACAT GG 32

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CGGGATCCGG GGTAATTTTT ACTAATTTCA TG 32

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CGGGATCCCG TATGTCTCAC TGTGTTTGTG GC 32

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

CGGGATCCCC CTACATACAT ATATCAGATT AG 32

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

CCATCGATAC TTTAAGTGAT TTGATGGC 28

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic DNA Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CGGGATCCTG TTCTGATTCT TGATTTCC 28

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2122 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

GTATGTCTCA CTGTGTTTGT GGCTGTGTGT GTTTTTTTCT CTGTCTTTTT GTGTTTTGTG 60

TAATTGGGGC TCTTTAAAGT TGGTATTGTG TATACCCTTT TGAGTATAGT CTTTGAGGAA 120

GCAAAATGAT GAATCTTGAT TGACATTAGT AAGGGTTGTA ACTTTTTGAA GTTTGGTTAG 180
GTGTAATTGA GTTTGGCTTG TGTGTCTGTG TGTCGAGGTT ATTTTTTTGG TTTGTGTTAT 240
TGGGGATTCT TAAAAGTTGG TATTGTGTAT ACCCTTTTGA GTATAGTCTT TGAGGAAGCA 300
AAAATGATGA ATCTTGATTG GCATTAGTAA AGGTTGTAGC TTTTTGAAGT GTGGTTAGGT 360
GTAATTGAGT TTGGCTTGTG TGTCTGTGTG TTTTGGAATC CTGATGTGTG TCAAGTCCTG 420
ATATGGGTCG AGGTTCTTTC TTTGGTTTGT GTAATTGGGG GTTCTTAAAA GTTGGTATTA 480
TGTACCTTTT TAAGAATAGT GTCTGAGAAA GCAAAATCGA TGAATTTTGA TTGACAGCAT 540
ATTCTTTGAG AAAGCAAAAA ATGGTGAGTT TTCATGGAGA AACTTGATTG ACATTACTAA 600
AGGTAGCAAC TTTTTCAACT CCTGATATGG GTCAAGGTTC TTTGTTTGGT TTGTGTAATT 660
TGGGGTTCTT TGAAGTTTTG AGAAAGAAAA ATTATGATTT TTCATGGAGA AATTTGATTT 720
ACATTAATAA AGGTAGTAGC TTTTTAAAGT GTGGTCAGCT GTAATGAGTT CAGCTTGGTT 780
TAAAGGGGCC CTACATATGG TGCTTTCTGG TGAGATATTT GTTGCTCCAC CATACGAGTT 840
ATAAGAATCA TAGTGTTAGG ATCTTTTTTC TTTTTTTTTT CATTTTTCAC TTGACTAGCT 900
ACTAGAGGAG TGATCTTGAC GGCGGAAAAT CTTAGAAAGG GGAAGGTTGT TTGCATCAAC 960
TGGTGTTATA TGTGCAAGGA GACGGGAGAT GATGTAGATC ATCTTCTTCT TCATTGTGGT 1020
CTTTCCATGA GGTTATGATG TGATATGTTT GAATGGTTTG GTACTTCTTG GCTATGCCAA 1080
GAACTGTGAA AGAATTGATA TTCAGTTGGA AGTGTGGAGT TGGAAGAGTG GAAGAATTGA 1140
CACTTGGTTC CATTAGCTTT AATGTGGGTG GTGTGGAGAG AGAGAGAAAT AGGAGAGCTT 1200
TTGAGGGGGT AGAGTTGAGC TTTCCTCAGT TGAGAAGTAG CCTTTGATAT CTTTTTTTTT 1260
TTTTTTTGTA CACCCATAGA ATTCCCAATT GTATAGAAGA TTGGGTGGAG TTTGTAGAGA 1320
ATCATCTTTT GTAGTAGATT CTTTACCTTT TGGTATATCC ATTGTATACA GCCAGGCCTT 1380
TGACTATGTT TATGAATGAA TATACATTAC TTGAAAAAAA AAGAAGTGAA GCCAGTCTGT 1440
TGTACCTTTG TAGACAATGT TGTTGCAGCA TCTTGATAAT TCCCTGAAAA TTGTCTCCCT 1500
GAAGGAATAG TTTGGTTGAT ATTGATTATT TCTTGGTTTG TTTAATTCGG TGTTCTTGAA 1560
GGCCATTTTA AATCCTTTGA CATTGTTAAA GGTGTTTACA AGTGTTGGTC TGGGTTTAAA 1620
AGCACCTCTT GTATGGTGCT TTCTGGAGTG ATCTTTCTTC CTCCAAAAGA GAAGTTGCAA 1680
GAATCAGTGT GTGTACTTTT TTCTCTTGTA TGATCAGATC TTTTTTCAAT TTTTCCGTTT 1740
TAGTTGATTT ATCCATATAG TGAAAGTTGG TGTCATAGTT GCTGTTTGTG GACTTCCTGT 1800
AAAAGTTTTT TGATATACTT AAAAAATTGT CACACAGAAG AAAGAGTTTT TTACCATTAC 1860

TTAAGCTAGA TGGGACTGTT TGATTCTTAG ACCAAATAAT GAACCTTTTT GTTCTCTTAA 1920

CGTGTACTTG AAATAGTTTG GTAAAATTGT GATAGGAAAA AAGATAATTC TTGATTGCTT 1980

TTGGAGCATC ACTTCTAATC ATAAAAGTCT TTGCTCTCTT CAACCATGAA TGATAAATTG 2040

GACACTTATG TGGCCCTAAG TTGCTCTCAG TAGTGGTCTT TAATTGTGGA GATATAACTA 2100

ATCTGATATA TGTATGTAGG GA 2122

CLAIMS

1. A method of affecting enzymatic activity in a plant (or a cell, a tissue or an organ thereof) comprising expressing in the plant (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron of a class A potato starch branching enzyme in an antisense orientation, optionally together with a nucleotide sequence which codes, partially or completely, for an intron of a class B starch branching enzyme in an antisense or sense orientation; and wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

2. A method according to claim 1 wherein starch branching enzyme activity is affected and/or wherein the levels of amylopectin are affected and/or the composition of starch is changed.

3. A method of affecting enzymatic activity in a starch producing organism (or a cell, a tissue or an organ thereof) comprising expressing in the starch producing organism (or a cell, a tissue or an organ thereof) a nucleotide sequence wherein the nucleotide sequence codes, partially or completely, for an intron of a class A starch branching enzyme in an antisense orientation, optionally together with a nucleotide sequence which codes, partially or completely, for an intron of a class B starch branching enzyme in an antisense or sense orientation; and wherein starch branching enzyme activity is affected and/or the levels of amylopectin are affected and/or the composition of starch is changed.

4. A method according to claim 3 wherein the nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

5. A method according to any one of the preceding claims wherein the enzymatic activity is reduced or eliminated.

6. A method according to any one of the preceding claims wherein the nucleotide sequence codes for at least substantially all of at least one intron in an antisense orientation.

7. A method according to any one of the preceding claims wherein the nucleotide sequence codes for all of at least one intron in an antisense orientation.

8. A method according to any one of the preceding claims wherein the nucleotide sequence comprises the complement of SEQ. ID. No. 38, or a fragment thereof.

9. A method according to any one of the preceding claims wherein the nucleotide sequence is expressed by a promoter having a sequence shown as SEQ.I.D. No. 14 or a variant, derivative or homologue thereof.

10. An antisense sequence comprising the nucleotide sequence as defined in claim 8 or a variant, derivative or homologue thereof.

11. A promoter having a sequence shown as SEQ.I.D. No. 14, or a variant, derivative or homologue thereof.

12. A promoter according to claim 11 in combination with a gene of interest ("GOI").

13. A construct capable of comprising or expressing the invention according to any one of claims 10 to 12.

14. A vector comprising or expressing the invention according to any one of claims 10 to 13.

15. A combination of nucleotide sequences comprising a first nucleotide sequence coding for a recombinant enzyme; and a second nucleotide sequence which corresponds to an intron in antisense orientation; wherein the intron is an intron that is associated with a genomic gene encoding an enzyme corresponding to the recombinant enzyme; and wherein the second nucleotide sequence does not contain a sequence that is antisense to an exon sequence normally associated with the intron.

16. A cell, tissue or organ comprising or expressing the invention according to any one of claims 10 to 15.

17. A transgenic starch producing organism comprising or expressing the invention according to any one of claims 10 to 16.

18. A transgenic starch producing organism according to claim 17 wherein the organism is a plant.

19. A starch obtained from the invention according to any one of the preceding claims.

20. A nucleotide sequence that is antisense to an intron of class A SBE.

21. A method for modifying starch production in an organism, comprising transforming the organism with a transgene capable of expressing an antisense intron sequence relating to class A SBE and a transgene capable of expressing an antisense intron sequence relating to class B SBE, thereby reducing or eliminating endogenous class A and class B production, and a further sequence encoding a SBE from a heterologous source.

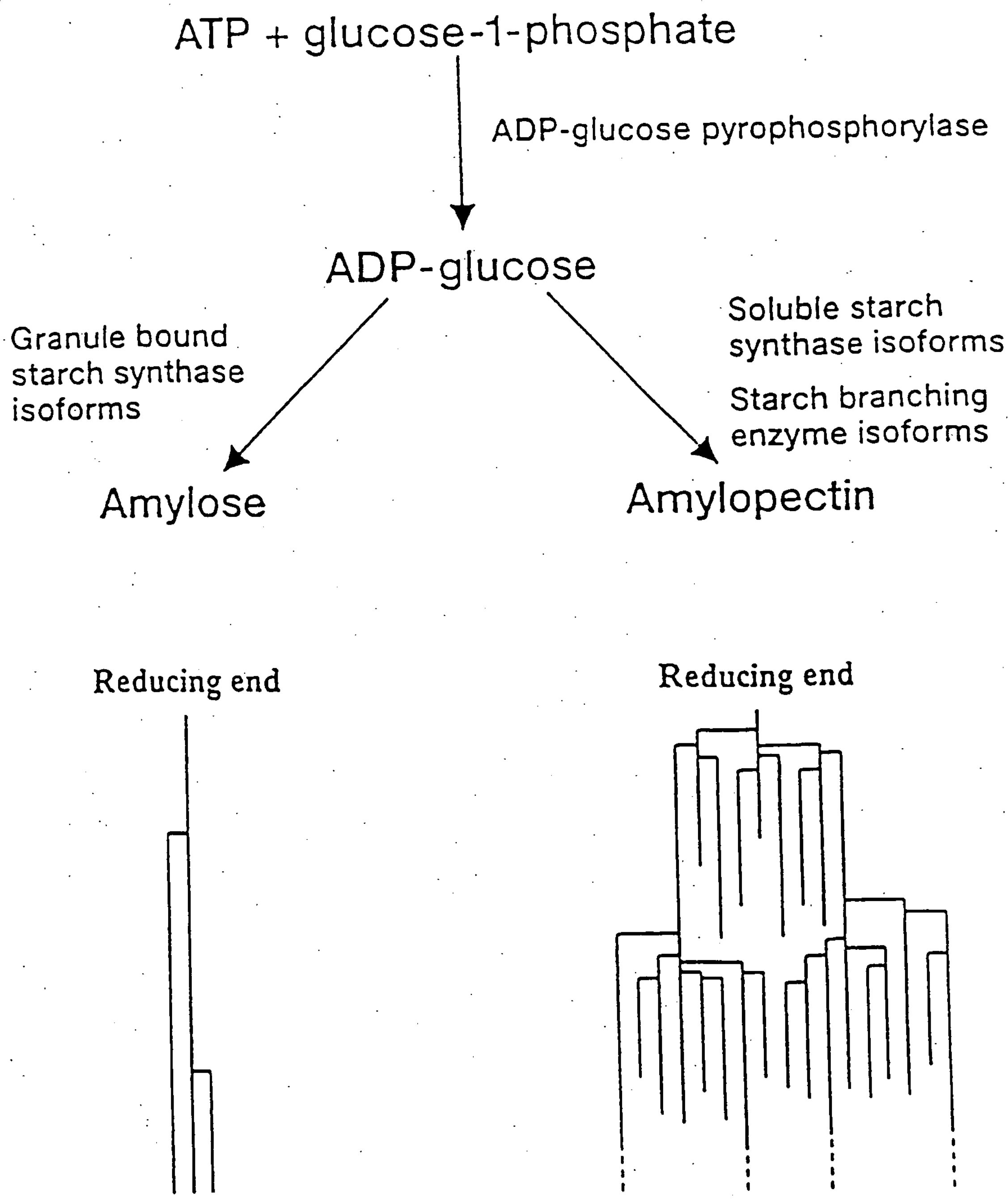

ATP + glucose-1-phosphate
ADP-glucose pyrophosphorylase
ADP-glucose
Granule bound starch synthase isoforms
Soluble starch synthase isoforms
Starch branching enzyme isoforms
Amylose
Amylopectin
Reducing end
Reducing end

FIG. 1

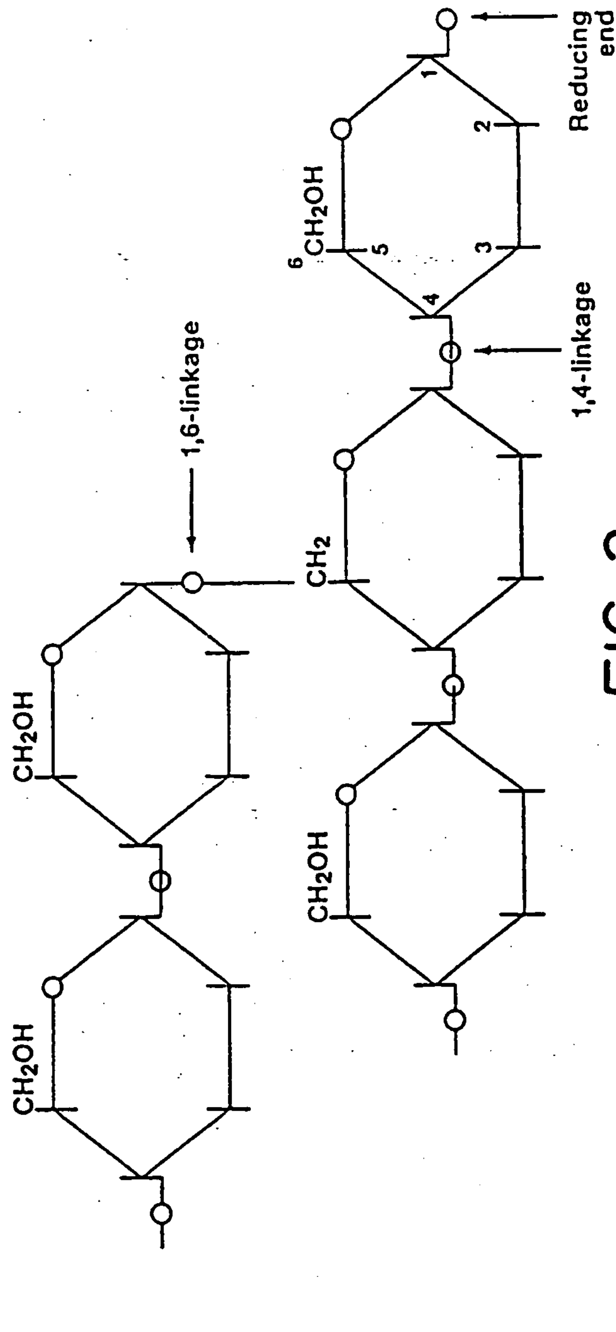

Reducing end
1,4-linkage
1,6-linkage
$^6CH_2OH$
1
2
3
4
5
CH_2
CH_2OH
CH_2OH
CH_2OH

FIG. 2

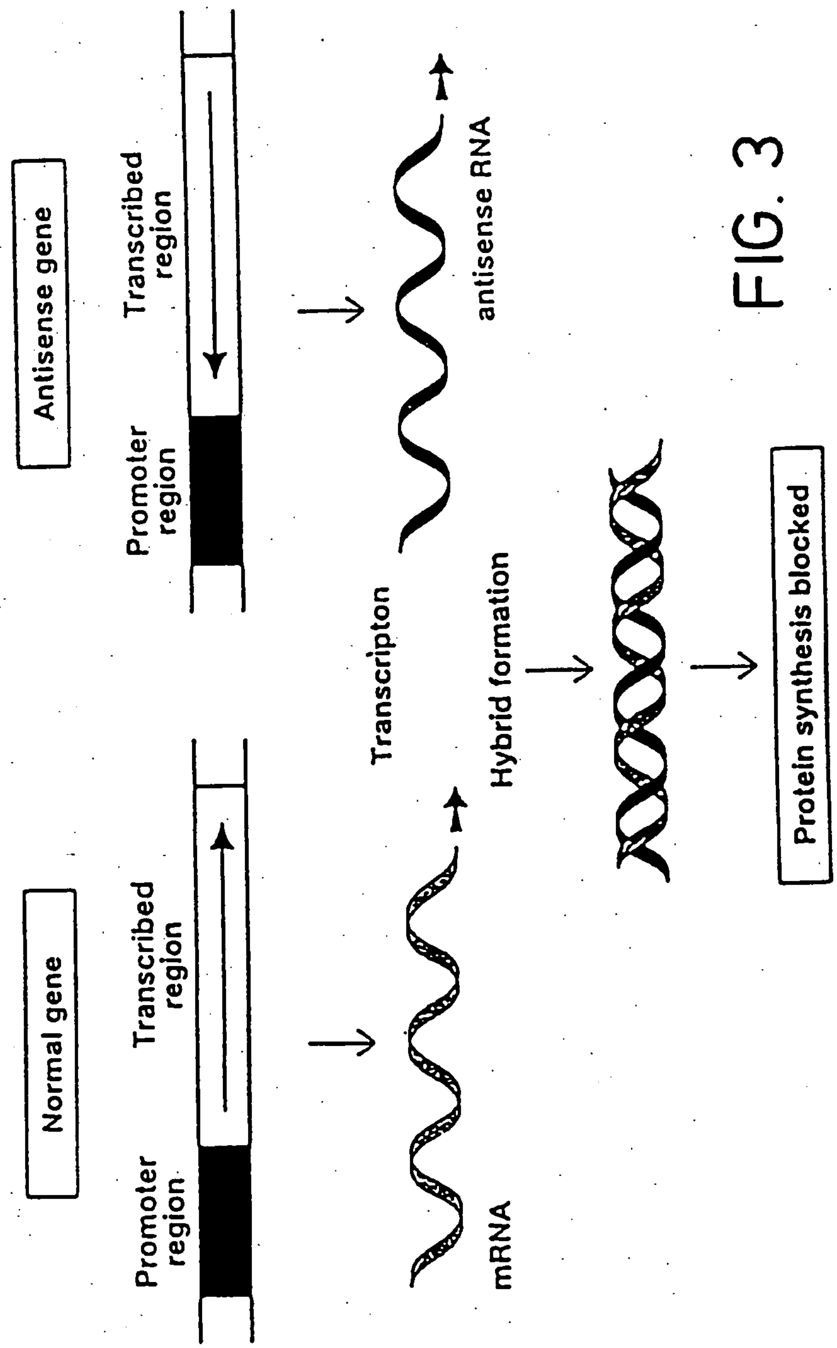

Antisense gene
Promoter region
Transcribed region
antisense RNA
Normal gene
Promoter region
Transcribed region
mRNA
Transcripton
Hybrid formation
Protein synthesis blocked

FIG. 3

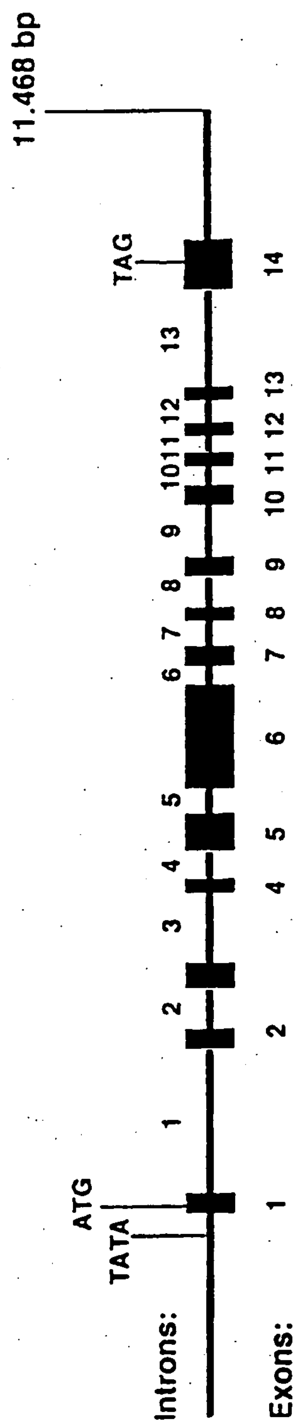

11.468 bp
TAG
ATG
TATA
Introns:
1
2
3
4
5
6
7
8
9
1011
12
13
Exons:
1
2
4
5
6
7
8
9
10
11
12
13
14

FIG. 4

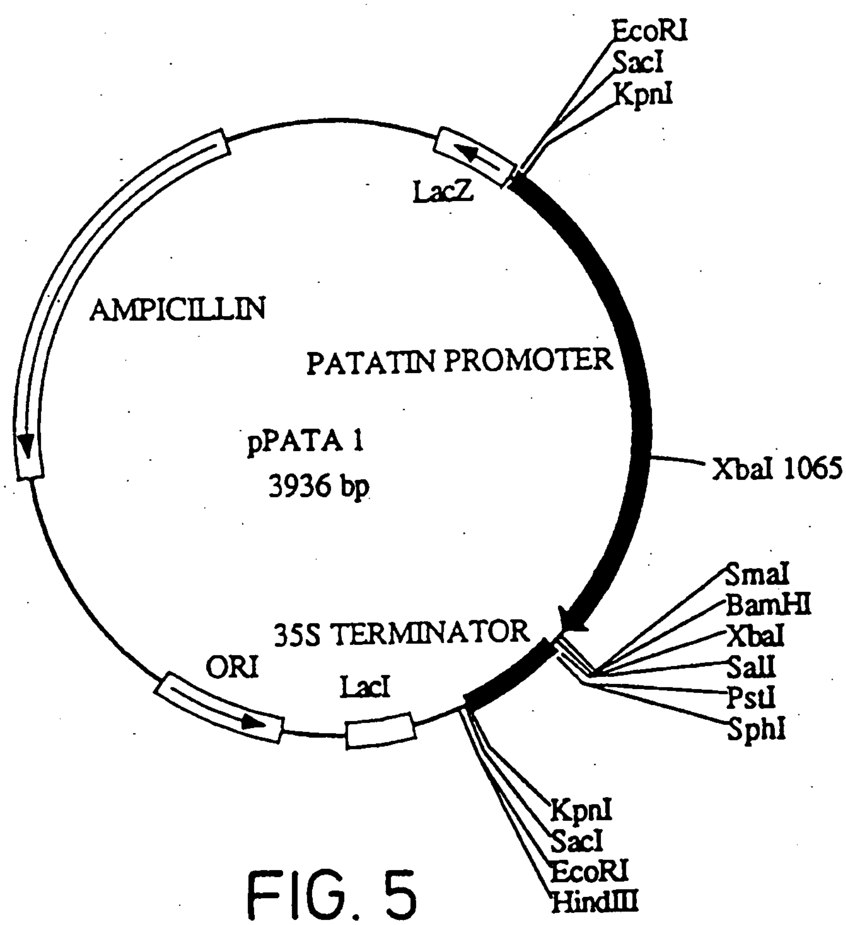

EcoRI
SacI
KpnI
LacZ
AMPICILLIN
PATATIN PROMOTER
pPATA 1
3936 bp
XbaI 1065
SmaI
BamHI
XbaI
SalI
PstI
SphI
35S TERMINATOR
ORI
LacI
KpnI
SacI
EcoRI
HindIII

FIG. 5

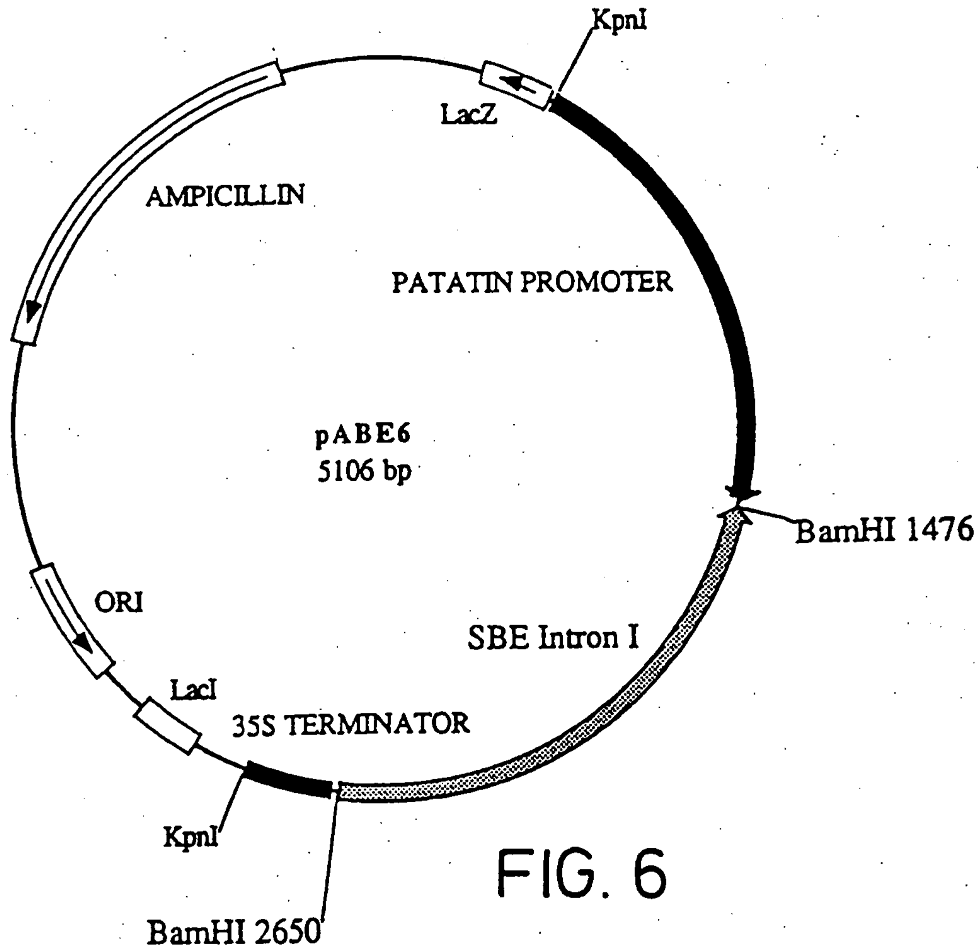

KpnI
LacZ
AMPICILLIN
PATATIN PROMOTER
pABE6
5106 bp
BamHI 1476
ORI
SBE Intron I
LacI
35S TERMINATOR
KpnI
BamHI 2650

FIG. 6

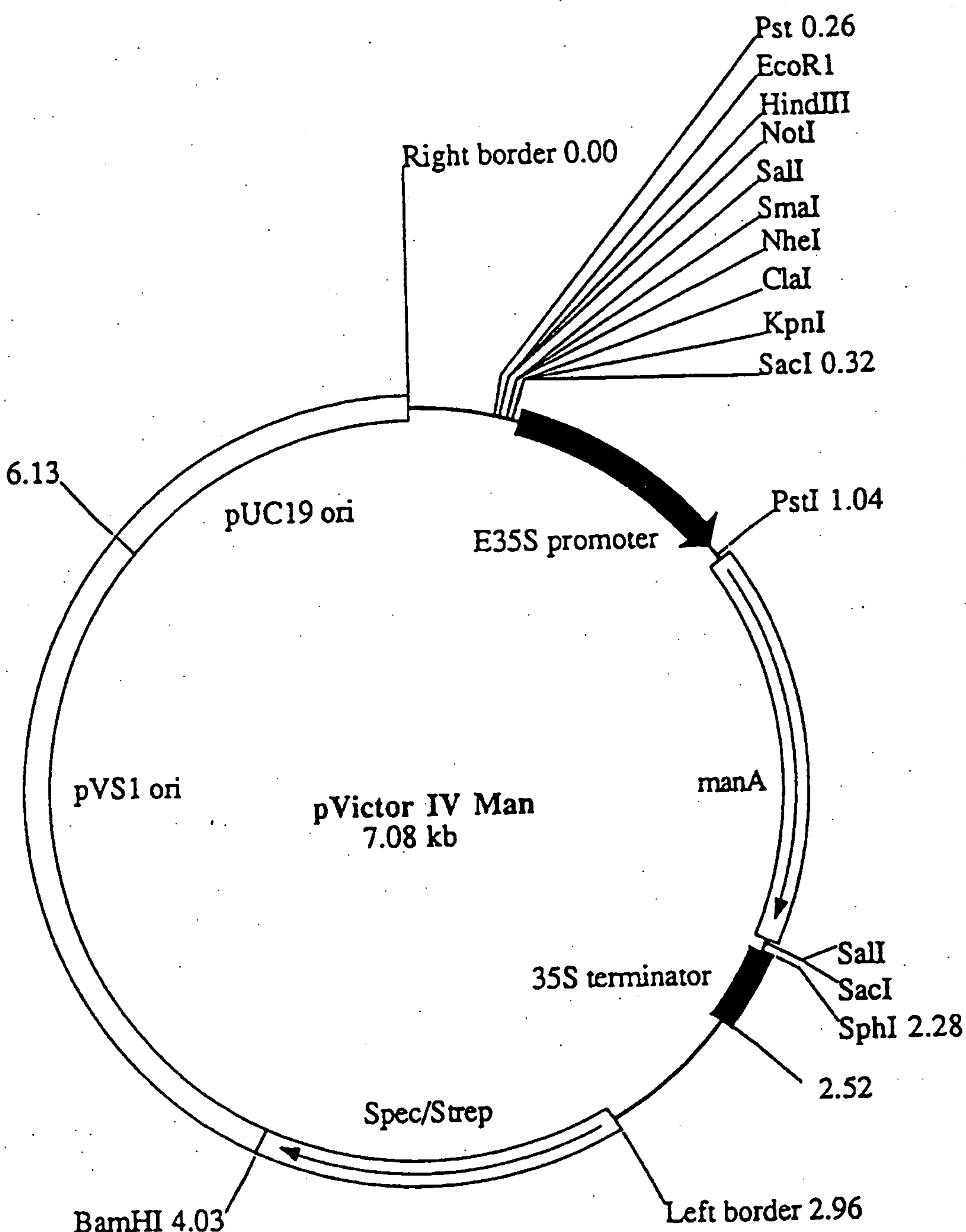

Pst 0.26
EcoR1
HindIII
NotI
SalI
SmaI
NheI
ClaI
KpnI
SacI 0.32
Right border 0.00
6.13
pUC19 ori
E35S promoter
PstI 1.04
manA
pVS1 ori
pVictor IV Man
7.08 kb
35S terminator
SalI
SacI
SphI 2.28
2.52
Spec/Strep
BamHI 4.03
Left border 2.96

FIG. 7

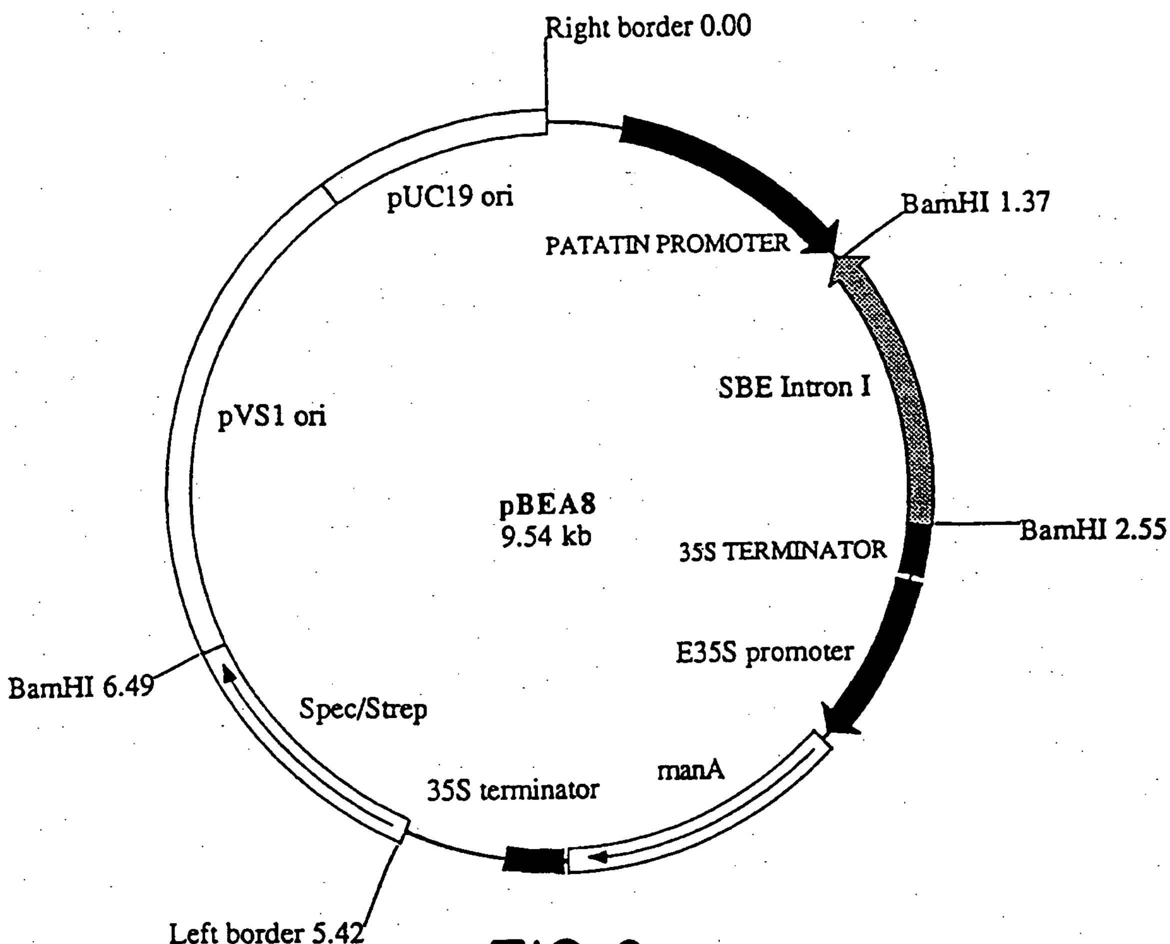

Right border 0.00
BamHI 1.37
PATATIN PROMOTER
pUC19 ori
SBE Intron I
pVS1 ori
pBEA8
9.54 kb
BamHI 2.55
35S TERMINATOR
E35S promoter
BamHI 6.49
Spec/Strep
manA
35S terminator
Left border 5.42

FIG. 8

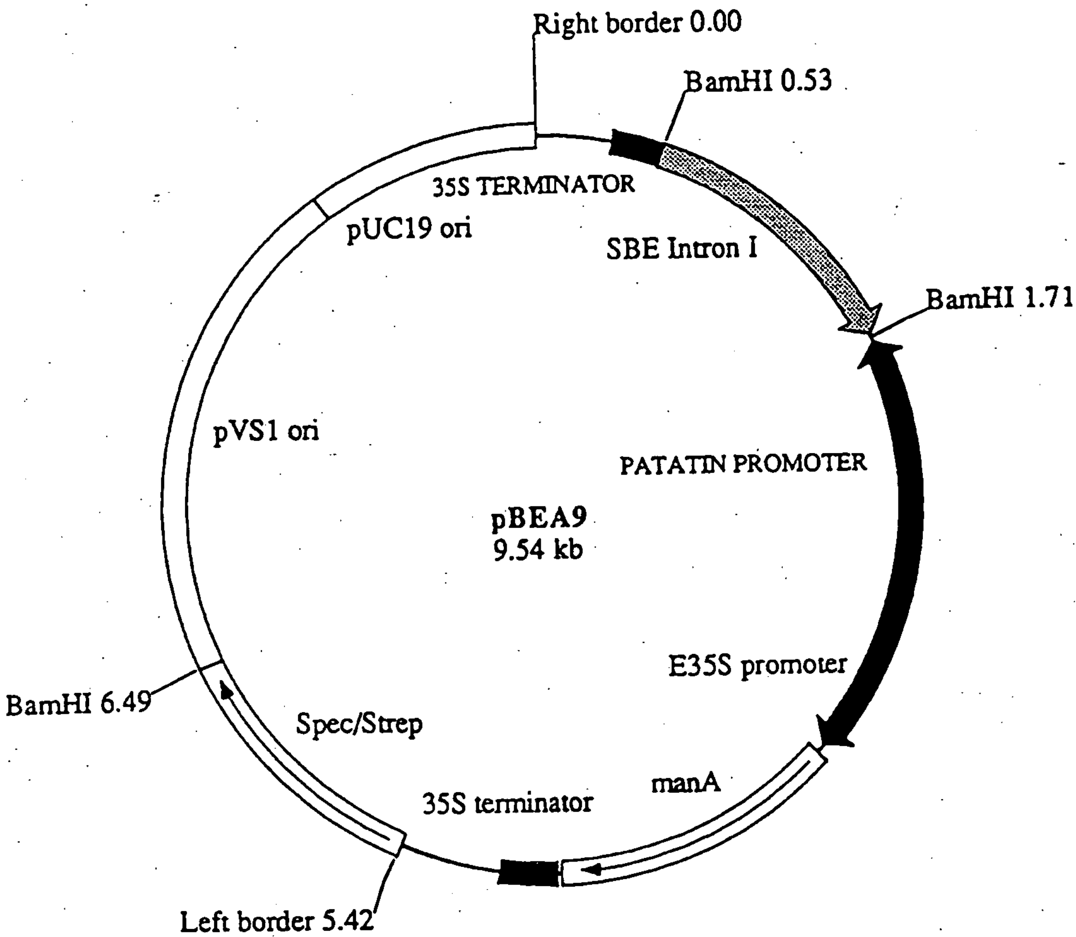

Right border 0.00
BamHI 0.53
35S TERMINATOR
pUC19 ori
SBE Intron I
BamHI 1.71
pVS1 ori
PATATIN PROMOTER
pBEA9
9.54 kb
E35S promoter
BamHI 6.49
Spec/Strep
manA
35S terminator
Left border 5.42

FIG. 9

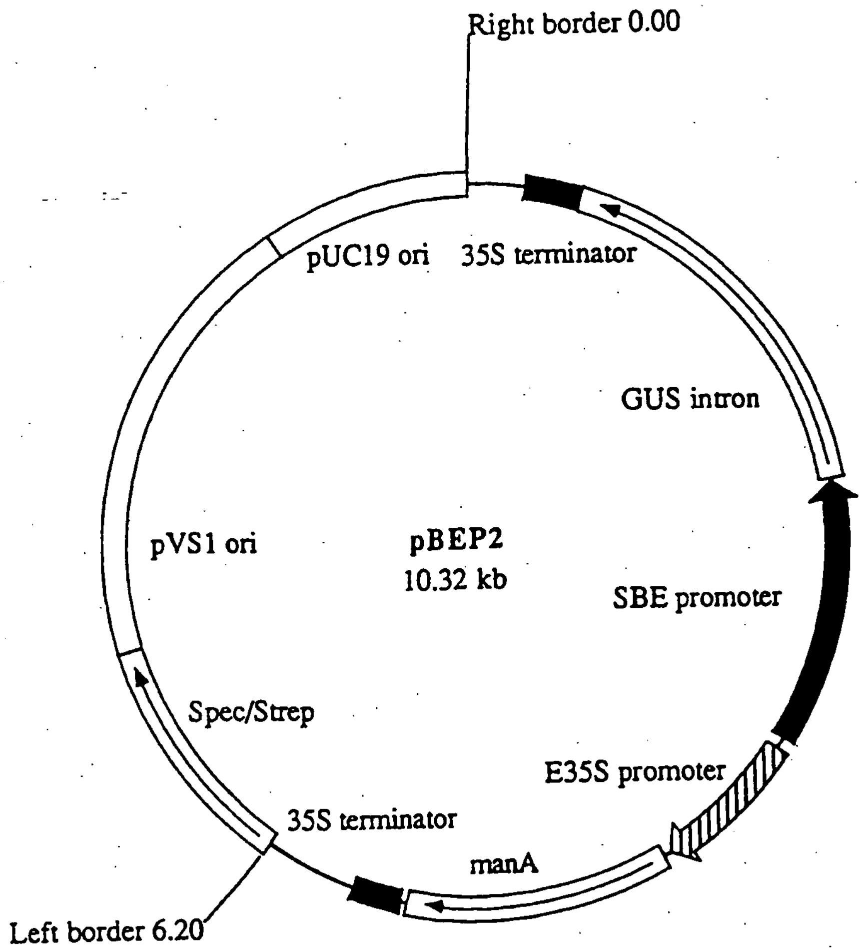

Right border 0.00
pUC19 ori
35S terminator
GUS intron
pVS1 ori
pBEP2
10.32 kb
SBE promoter
Spec/Strep
E35S promoter
35S terminator
manA
Left border 6.20

FIG. 10

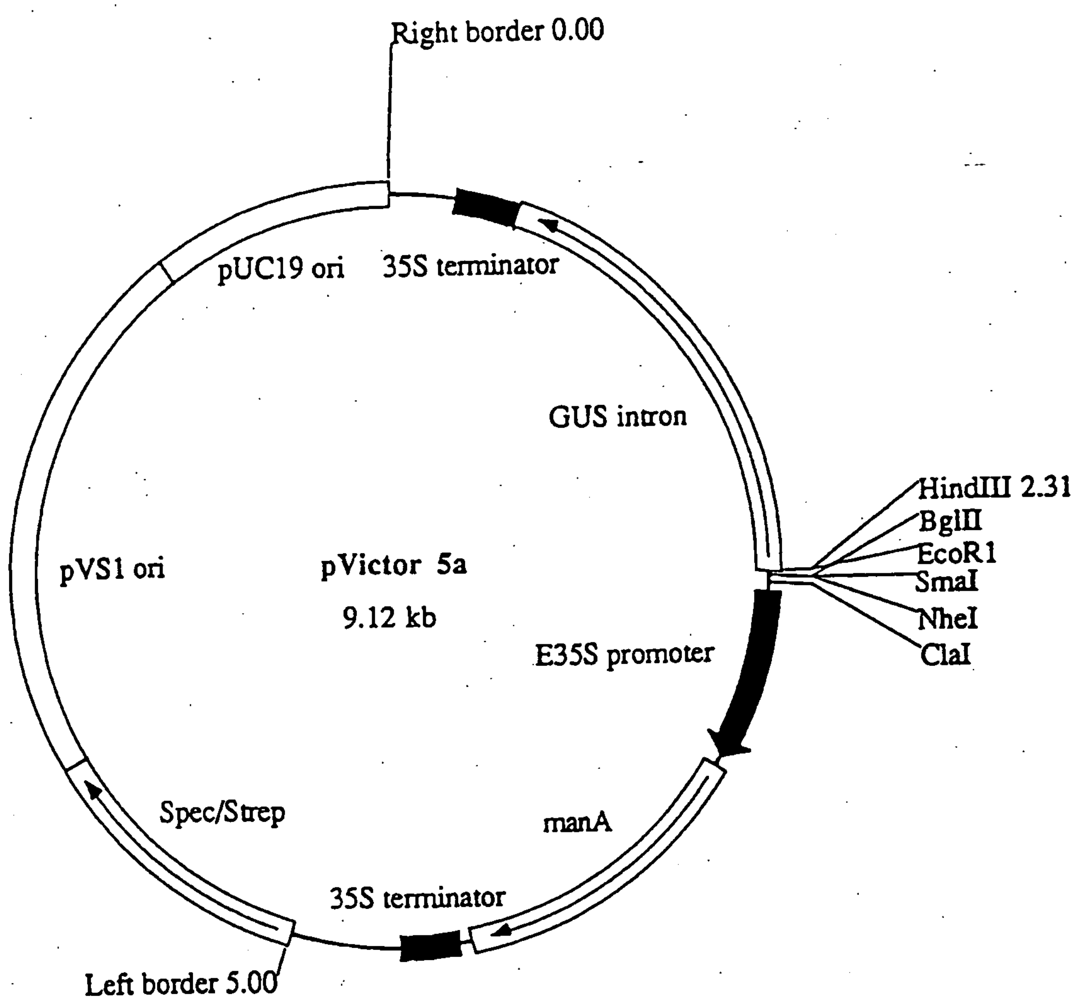

Right border 0.00
pUC19 ori
35S terminator
GUS intron
HindIII 2.31
BglII
EcoR1
SmaI
NheI
ClaI
pVS1 ori
pVictor 5a
9.12 kb
E35S promoter
Spec/Strep
manA
35S terminator
Left border 5.00

FIG. 11

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
ATCATGGCCAATTACTGGTTCAAATGCATTACTTCCTTTCAGATTCTTTCGAGTTCTCAT     60
GACCGGTCCTACTACAGACGATACTAACCCGTGGAACTGTTGCATCTGCTTCTTAGAACT    120
CTATGGCTATTTTCGTTAGCTTGGCGTCGGTTTGAACATAGTTTTTGTTTTCAAACTCTT    180
CATTTACAGTCAAAATGTTGTATGGTTTTTGTTTTCCTCAATGATGTTTACAGTGTTGTG    240
TTGTCATCTGTACTTTTGCCTATTACTTGTTTTGAGTTACATGTTAAAAAAGTGTTTATT    300
TTGCCATATTTTGTTCTCTTATTATTATTATCATACATACATTATTACAAGGAAAAGACA    360
AGTACACAGATCTTAACGTTTATGTTCAATCAACTTTTGGAGGCATTGACAGGTACCACA    420
AATTTTGAGTTTATGATTAAGTTCAATCTTAGAATATGAATTTAACATCTATTATAGATG    480
CATAAAAATAGCTAATGATAGAACATTGACATTTGGCAGAGCTTAGGGTATGGTATATCC    540
AACGTTAATTTAGTAATTTTTGTTACGTACGTATATGAAATATTGAATTAATCACATGAA    600
CGGTGGATATTATATTATGAGTTGGCATCAGCAAAATCATTGGTGTAGTTGACTGTAGTT    660
GCAGATTTAATAATAAAATGGTAATTAACGGTCGATATTAAAATAACTCTCATTTCAAGT    720
GGGATTAGAACTAGTTATTAAAAAAATGTATACTTTAAGTGATTTGATGGCATATAATTT    780
AAAGTTTTTCATTTCATGCTAAAATTGTTAATTATTGTAATGTAGACTGCGACTGGAATT    840
ATTATAGTGTAAATTTATGCATTCAGTGTAAAATTAAAGTATTGAACTTGTCTGTTTTAG    900
AAAATACTTTATACTTTAATATAGGATTTTGTCATGCGAATTTAAATTAATCGATATTGA    960
ACACGGAATACCAAAATTAAAAAGGATACACATGGCCTTCATATGAACCGTGAACCTTTG   1020
ATAACGTGGAAGTTCAAAGAAGGTAAAGTTTAAGAATAAACTGACAAATTAATTTCTTTT   1080
ATTTGGCCCACTACTAAATTTGCTTTACTTTCTAACATGTCAAGTTGTGCCCTCTTAGTT   1140
GAATGATATTCATTTTTCATCCCATAAGTTCAATTTGATTGTCATACCACCCATGATGTT   1200
CTGAAAAATGCTTGGCCATTCACAAAGTTTATCTTAGTTCCTATGAACTTTATAAGAAGC   1260
TTTAATTTGACATGTTATTTATATTAGATGATATAATCCATGACCCAATAGACAAGTGTA   1320
TTAATATTGTAACTTTGTAATTGAGTGTGTCTACATCTTATTCAATCATTTAAGGTCATT   1380
AAAATAAATTATTTTTTGACATTCTAAAACTTTAAGCAGAATAAATAGTTTATCAATTAT   1440
TAAAAACAAAAAACGACTTATTTATAAATCAACAAACAATTTTAGATTGCTCCAACATAT   1500
```

FIG. 12

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
TTTTCCAAATTAAATGCAGAAAATGCATAATTTTATACTTGATCTTTATAGCTTATTTTT   1560

TTTAGCCTAACCAACGAATATTTGTAAACTCACAACTTGATTAAAAGGGATTTACAACAA   1620

GATATATATAAGTAGTGACAAATCTTGATTTTAAATATTTTAATTTGGAGGTCAAAATTT   1680

TACCATAATCATTTGTATTTATAATTAAATTTTAAATATCTTATTTATACATATCTAGTA   1740

AACTTTTAAATATACGTATATACAAAATATAAAATTATTGGCGTTCATATTAGGTCAATA   1800

AATCCTTAACTATATCTGCCTTACCACTAGGAGAAAGTAAAAAACTCTTTACCAAAAATA   1860

CATGTATTATGTATACAAAAAGTCGATTAGATTACCTAAATAGAAATTGTATAACGAGTA   1920

AGTAAGTAGAAATATAAAAAAACTACAATACTAAAAAAAATATGTTTTACTTCAATTTCG   1980

AAACTAATGGGGTCTGAGTGAAATATTCAGAAAGGGGAGGACTAACAAAAGGGTCATAAT   2040

GTTTTTTTATAAAAAGCCACTAAAATGAGGAAATCAAGAATCAGAACATACAAGAAGGCA   2100

GCAGCTGAAGCAAAGTACCATAATTTAATCAATGGAAATTAATTTCAAAGTTTTATCAAA   2160
                                  M  E  I  N  F  K  V  L  S  K
ACCCATTCGAGGATCTTTTCCATCTTTCTCACCTAAAGTTTCTTCAGGGgtaattttac   2220
 P  I  R  G  S  F  P  S  F  S  P  K  V  S  S  G
taatttcatgttaatttcaattatttttagcctttgcatttcattttccaatatatctgg   2280

atcatctccttagtttttttattttatttttttataatatcaaatatggaagaaaaatgaca   2340

cttgtagagccatatgtaagtatcatgtgacaaatttgcaaggtggttgagtgtataaaa   2400

ttcaaaaattgagagatggaggggggtgggggbaragacaatatttagaaagagtgttc   2460

taggaggttatggaggacacggatgaggggtagaaggttagttaggtatttgagtgttgt   2520

ctggcttatcctttcatactagtagtcgtggaattatttgggtagtttcttgttttgtta   2580

tttgatctttgttattctattttctgtttcttgtacttcgattattgtattatatatctt   2640

gtcgtagttattgttcctcggtaagaatgctctagcatgcttcctttagtgttttatcat   2700

gccttctttatattcgcgttgctttgaaatgcttttactttagccgagggtctattagaa   2760

acaatctctctatctcgtaaggtaggggtaaagtcctcaccacactccacttgtgggatt   2820

acattgtgtttgttgttgtaaatcaattatgtatacataataagtggatttttacaaca   2880

caaatacatggtcaagggcaaagttctgaacacataaagggttcattatatgtccaggga   2940

tatgataaaaattgtttctttgtgaaagttatataagatttgttatggcttttgctggaa   3000
```

FIG. 12 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
acataataagttataatgctgagatagctactgaagtttgtttttctagccttttaaat   3060

gtaccaataatagattccgtatcgaacgagtatgttttgattacctggtcatgatgtttc  3120

tatttttacattttttggtgttgaactgcaattgaaaatgttgtatcctatgagacgg    3180

atagttgagaatgtgttctttgtatggaccttgagaagctcaaacgctactccaataatt  3240

tctatgaattcaaattcagtttatggctaccagtcagtccagaaattaggatatgctgca  3300

tatacttgttcaattatactgtaaaatttcttaagttctcaagatatccatgtaacctcg  3360

agaatttctttgacagGCTTCTAGAAATAAGATATGTTTTCCTTCTCAACATAGTACTGG  3420
                 A  S  R  N  K  I  C  F  P  S  Q  H  S  T  G
ACTGAAGTTTGGATCTCAGGAACGGTCTTGGGATATTTCTTCCACCCCAAAATCAAGAGT  3480
 L  K  F  G  S  Q  E  R  S  W  D  I  S  S  T  P  K  S  R  V
TAGAAAAGATGAAAGGgtatgtttgataatttatatggttgcatggatagtatataaata  3540
 R  K  D  E  R
gttggaaaacttctggactggtgctcatggcatatttgatctgtgcaccgtgtggagatg  3600

tcaaacatgtgttacttcgttccgccaatttataataccttaacttgggaaagacagctc  3660

tttactcctgtgggcatttgttatttgaattacaatctttatgagcatggtgttttcaca  3720

ttatcaacttctttcatgtggtatataacagtttttagctccgttaatacctttcttctt  3780

tttgatataaactaactgtggtgcattgcttgcbkkkATGAAGCACAGTTCAGCTATTTC  3840
                                       M  K  H  S  S  A  I  S
CGCTGTTTTGACCGATGACGACAATTCGACAATGGCACCCCTAGAGGAAGATGTCAAGAC  3900
 A  V  L  T  D  D  D  N  S  T  M  A  P  L  E  E  D  V  K  T
TGAAAATATTGGCCTCCTAAATTTGGATCCAACTTTGGAACCTTATCTAGATCACTTCAG  3960
 E  N  I  G  L  L  N  L  D  P  T  L  E  P  Y  L  D  H  F  R
ACACAGAATGAAGAGATATGTGGATCAGAAAATGCTCATTGAAAAATATGAGGGACCCCT  4020
 H  R  M  K  R  Y  V  D  Q  K  M  L  I  E  K  Y  E  G  P  L
TGAGGAATTTGCTCAAGgtaacagccaaaagttgtgctttaggcagtttgaccttatttt  4080
 E  E  F  A  Q  G
ggaagatgaattgtttatacctactttgactttgctagagaattttgcataccggggagt  4140

aagtagtggctccatttaggtggcacctggccatttttttgatcttttaaaaagctgttt  4200

gattgggtcttcaaaaaagtagacaaggtttttggagaagtgacacacccccggagtgtc  4260

agtggcaaagcaaagattttcactaaggagattcaaaatataaaaaaagtatagacataa  4320

agaagctgaggggattcaacatgtactatacaagcatcaaatatagtcttaaagcaattt  4380

tgtagaaataaagaaagtcttccttctgttgcttcacaatttccttctattatcatgagt  4440

tactctttctgttcgaaatagcttccttaatattaaattcatgatacttttgttgagatt  4500
```

FIG. 12 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
tagcagtttttcttgtgtaaactgctctctttttttgcagGTTATTTAAAATTTGGATT   4560
                                          Y  L  K  F  G  F
CAACAGGGAAGATGGTTGCATAGTCTATCGTGAATGGGCTCCTGCTGCTCAgtaggtcct  4620
 N  R  E  D  G  C  I  V  Y  R  E  W  A  P  A  A  Q
cgtctactacaaaatagtagtttccatcatcataacagattttcctattaaagcatgatg  4680

ttgcagcatcattggctttcttacatgttctaattgctattaaggttatgcttctaatta  4740

actcatccacaatgcagGGAAGCAGAAGTTATTGGCGATTTCAATGGATGGAACGGTTCT  4800
                   E  A  E  V  I  G  D  F  N  G  W  N  G  S
AACCACATGATGGAGAAGGACCAGTTTGGTGTTTGGAGTATTAGAATTCCTGATGTTGAC  4860
N  H  M  M  E  K  D  Q  F  G  V  W  S  I  R  I  P  D  V  D
AGTAAGCCAGTCATTCCACACAACTCCAGAGTTAAGTTTCGTTTCAAACATGGTAATGGA  4920
S  K  P  V  I  P  H  N  S  R  V  K  F  R  F  K  H  G  N  G
GTGTGGGTAGATCGTATCCCTGCTTGGATAAAGTATGCCACTGCAGACGCCACAAAGTTT  4980
V  W  V  D  R  I  P  A  W  I  K  Y  A  T  A  D  A  T  K  F
GCAGCACCATATGATGGTGTCTACTGGGACCCACCACCTTCAGAAAGgttttgttattca  5040
A  A  P  Y  D  G  V  Y  W  D  P  P  P  S  E  R
taccttgaagctgaattttgaacaccatcatcacaggcatttcgattcatgttcttacta  5100

gtcttgttatgtaagacattttgaaatgcaaaagttaaaataattgtgtctttactaatt  5160

tggacttgatcccatactctttcccttaacaaaatgagtcaattctataagtgcttgaga  5220

acttactacttcagcaattaaacagGTACCACTTCAAATACCCTCGCCCTCCCAAACCCC  5280
                           Y  H  F  K  Y  P  R  P  P  K  P  R
GAGCCCCACGAATCTATGAAGCACATGTCGGCATGAGCAGCTCTGAGCCACGTGTAAATT  5340
  A  P  R  I  Y  E  A  H  V  G  M  S  S  S  E  P  R  V  N  S
CGTATCGTGAGTTTGCAGATGATGTTTTACCTCGGATTAAGGCAAATAACTATAATACTG  5400
  Y  R  E  F  A  D  D  V  L  P  R  I  K  A  N  N  Y  N  T  V
TCCAGTTGATGGCCATAATGGAACATTCTTACTATGGATCATTTGGATATCATGTTACAA  5460
  Q  L  M  A  I  M  E  H  S  Y  Y  G  S  F  G  Y  H  V  T  N
ACTTTTTTGCTGTGAGCAGTAGATATGGAAACCCGGAGGACCTAAAGTATCTGATAGATA  5520
  F  F  A  V  S  S  R  Y  G  N  P  E  D  L  K  Y  L  I  D  K
AAGCACATAGCTTGGGTTTACAGGTTCTGGTGGATGTAGTTCACAGTCATGCAAGCAATA  5580
  A  H  S  L  G  L  Q  V  L  V  D  V  V  H  S  H  A  S  N  N
ATGTCACTGATGGCCTCAATGGCTTTGATATTGGCCAAGGTTCTCAAGAATCCTACTTTC  5640
  V  T  D  G  L  N  G  F  D  I  G  Q  G  S  Q  E  S  Y  F  H
ATGCTGGAGAGCGAGGGTACCATAAGTTGTGGGATAGCAGGCTGTTCAACTATGCCAATT  5700
  A  G  E  R  G  Y  H  K  L  W  D  S  R  L  F  N  Y  A  N  W
GGGAGGTTCTTCGTTTCCTTCTTTCCAACTTGAGGTGGTGGCTAGAAGAGTATAACTTTG  5760
  E  V  L  R  F  L  L  S  N  L  R  W  W  L  E  E  Y  N  F  D
ACGGATTTCGATTTGATGGAATAACTTCTATGCTGTATGTTCATCATGGAATCAATATGG  5820
  G  F  R  F  D  G  I  T  S  M  L  Y  V  H  H  G  I  N  M  G
GATTTACAGGAAACTATAATGAGTATTTCAGCGAGGCTACAGATGTTGATGCTGTGGTCT  5880
  F  T  G  N  Y  N  E  Y  F  S  E  A  T  D  V  D  A  V  V  Y
ATTTAATGTTGGCCAATAATCTGATTCACAAGATTTTCCCAGATGCAACTGTTATTGCCG  5940
  L  M  L  A  N  N  L  I  H  K  I  F  P  D  A  T  V  I  A  E
AAGATGTTTCTGGTATGCCGGGCCTTGGCCGGCCTGTTTCTGAGGGAGGAATTGGTTTTG  6000
  D  V  S  G  M  P  G  L  G  R  P  V  S  E  G  G  I  G  F  V
```

FIG. 12 CONTINUED

```
          10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
TTTACCGCCTGGCAATGGCAATCCCAGATAAGTGGATAGATTATTTAAAGAATAAGAATG   6060
 Y  R  L  A  M  A  I  P  D  K  W  I  D  Y  L  K  N  K  N  D
ATGAAGATTGGTCCATGAAGGAAGTAACATCGAGTTTGACAAATAGGAGATATACAGAGA   6120
 E  D  W  S  M  K  E  V  T  S  S  L  T  N  R  R  Y  T  E  K
AGTGTATAGCATATGCGGAGACCCATGATCAGgtattttaaatttatttctacaactaaa   6180
 C  I  A  Y  A  E  T  H  D  Q
taattctcagaacaattgttagatagaatccaaatatatacgtcctgaaagtataaaagt   6240

acttattttcgccatgggccttcagaatattggtagccgctgaatatcatgataagttat   6300

ttatccagtgacatttttatgttcactcctattatgtctgctggatacagTCTATTGTTG   6360
                                                   S  I  V  G
GTGACAAGACCATTGCATTTCTCCTAATGGACAAAGAGATGTATTCTGGCATGTCTTGCT   6420
 D  K  T  I  A  F  L  L  M  D  K  E  M  Y  S  G  M  S  C  L
TGACAGATGCTTCTCCTGTTGTTGATCGAGGAATTGCGCTTCACAAGgtttgtctgtttc   6480
 T  D  A  S  P  V  V  D  R  G  I  A  L  H  K
tattgcattttaaggttcatataggttagccacggaaaatctcactctttgtgaggtaac   6540

cagggttctgatggattattcaattttctcgtttatcatttgtttattcttttcatgcat   6600

tgtgtttctttttcaatatccctcttatttggaggtaatttttctcatctattcactttt   6660

agcttctaaccacagATGATCCATTTTTTCACAATGGCCTTGGGAGGAGAGGGGTACCTC   6720
               M  I  H  F  F  T  M  A  L  G  G  E  G  Y  L
AATTTCATGGGTAACGAGgtatgtcttacatctttagatattttgtgataattacaatta   6780
N  F  M  G  N  E
gtttggcttacttgaacaagattcattcctcaaaatgacctgaactgttgaacatcaaag   6840

gggttgaaacatagaggaaaacaacatgatgaatgtttccattgtctagggatttctatt   6900

atgttgctgagaacaaatgtcatcttaaaaaaacattgtttactttttgtagtatataga   6960

agattactgtatagagtttgcaagtgtgtctgttttggagtaattgtgaaatgtttgatg   7020

aacttgtacagTTTGGCCATCCTGAGTGGATTGACTTCCCTAGAGAGGGCAATAATTGGA   7080
           F  G  H  P  E  W  I  D  F  P  R  E  G  N  N  W  S
GTTATGACAAATGTAGACGCCAGTGGAACCTCGCGGATAGCGAACACTTGAGATACAAGg   7140
  Y  D  K  C  R  R  Q  W  N  L  A  D  S  E  H  L  R  Y  K
ttcaagtattttgaatcgcagcttgttaaataatctagtaattttagattgcttacttg   7200

gaagtctacttggttctggggatgatagctcatttcatcttgttctacttattttccaac   7260

cgaatttctgattttgtttcgagatccaagtattagattcatttacacttattaccgcc   7320

tcatttctaccactaaggccttgatgagcagcttaagttgattctttgaagctatagttt   7380

caggctaccaatccacagcctgctatatttgttggatacttaccttttctttacaatgaa   7440

gtgatactaattgaaatggtctaaatctgatatctatatttctccgtctttcctcccccct   7500
```

FIG. 12 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
catgatgaaatgcagTTTATGAATGCATTTGATAGAGCTATGAATTCGCTCGATGAAAAG  7560
               F  M  N  A  F  D  R  A  M  N  S  L  D  E  K
TTCTCATTCCTCGCATCAGGAAAACAGATAGTAAGCAGCATGGATGATGATAATAAGgta  7620
F  S  F  L  A  S  G  K  Q  I  V  S  S  M  D  D  D  N  K
aaatcatctaaagttgaaagtgttgggtttatgaagtgctttaattctatccaaggacaa  7680

gtagaaacctttttaccttccatttcttgatgatggatttcatattatttaatccaatag  7740

ctggtcaaattcggtaatagctgtactgattagttacttcactttgcagGTTGTTGTGTT  7800
                                                  V  V  V  F
TGAACGTGGTGACCTGGTATTTGTATTCAACTTCCACCCAAAGAACACATACGAAGGgta  7860
 E  R  G  D  L  V  F  V  F  N  F  H  P  K  N  T  Y  E  G
tatatgttttacttatccatgaaattattgctctgcttgttttaatgtactgaacaagt  7920

tttatggagaagtaactgaaacaaatcattttcacattgtctaatttaactcttttttct  7980

gatcctcgcatgacgaaaacagGTATAAAGTTGGATGTGACTTGCCAGGGAAGTACAGAG  8040
                       Y  K  V  G  C  D  L  P  G  K  Y  R  V
TTGCACTGGACAGTGATGCTTGGGAATTTGGTGGCCATGGAAGAgtaaggatttgcttga  8100
  A  L  D  S  D  A  W  E  F  G  G  H  G  R
ataactttgataataagataacagatgtagggtacagttctctcaccaaaaagaactgt  8160

aattgtctcatccatctttagttgtataagatatccgactgtctgagttcggaagtgttt  8220

gagcctcctgccctccccctgcgttgtttagctaattcaaaaaggagaaaactgtttatt  8280

gatgatctttgtcttcatgctgacatacaatctgttctcatgacagACTGGTCATGATGT  8340
                                                T  G  H  D  V
TGACCATTTCACATCACCAGAAGGAATACCTGGAGTTCCAGAAACAAATTTCAATGGTCG  8400
 D  H  F  T  S  P  E  G  I  P  G  V  P  E  T  N  F  N  G  R
TCCAAATTCCTTCAAAGTGCTGTCTCCTGCGCGAACATGTGTGgtacagttcttgccgtg  8460
 P  N  S  F  K  V  L  S  P  A  R  T  C  V
tgacctcccttttattgtggttttgttcatagttatttgaatgcgatagaagttaacta  8520

ttgattaccgccacaatcgccagttaagtcctctgaactactaatttgaaaggtaggaat  8580

agccgtaataaggtctacttttggcatcttactgttacaaaacaaaaggatgccaaaaaa  8640

attcttctctatcctctttttccctaaaccagtgcatgtagcttgcacctgcataaactt  8700

aggtaaatgatcaaaaatgaagttgatgggaacttaaaaccgccctgaagtaaagctagg  8760

aatagtcatataatgtccacctttggtgtctgcgctaacatcaacaacaacatacctcgt  8820

gtagtcccacaaagtggtttcagggggagggtagagtgtatgcaaaacttactcctatct  8880

cagaggtagagaggatttttcaatagacccttggctcaagaaaaaaagtccaaaaagaa  8940

gtaacagaagtgaaagcaacatgtgtagctaaagcgacccaacttgtttgggactgaagt  9000
```

FIG. 12 CONTINUED

```
          10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
agttgttgttgttgaaacagtgcatgtagatgaacacatgtcagaaaatggacaacacag   9060

ttattttgtgcaagtcaaaaaaatgtactactatttctttgtgcagctttatgtatagaa   9120

aagttaaataactaatgaattttgctagcagaaaaatagcttggagagaaatttttttata   9180

ttgaactaagctaactatattcatctttctttttgcttcttcttctccttgtttgtgaag   9240

GCTTATTACAGAGTTGATGAACGCATGTCAGAAACTGAAGATTACCAGACAGACATTTGT   9300
 A  Y  Y  R  V  D  E  R  M  S  E  T  E  D  Y  Q  T  D  I  C
AGTGAGCTACTACCAACAGCCAATATCGAGGAGAGTGACGAGAAACTTAAAGATTCGTTA   9360
 S  E  L  L  P  T  A  N  I  E  E  S  D  E  K  L  K  D  S  L
TCTACAAATATCAGTAACATTGACGAACGCATGTCAGAAACTGAAGTTTACCAGACAGAC   9420
 S  T  N  I  S  N  I  D  E  R  M  S  E  T  E  V  Y  Q  T  D
ATTTCTAGTGAGCTACTACCAACAGCCAATATTGAGGAGAGTGACGAGAAACTTAAAGAT   9480
 I  S  S  E  L  L  P  T  A  N  I  E  E  S  D  E  K  L  K  D
TCGTTATCTACAAATATCAGTAACATTGATCAGACTGTTGTAGTTTCTGTTGAGGAGAGA   9540
 S  L  S  T  N  I  S  N  I  D  Q  T  V  V  V  S  V  E  E  R
GACAAGGAACTTAAAGATTCACCGTCTGTAAGCATCATTAGTGATGTTGTTCCAGCTGAA   9600
 D  K  E  L  K  D  S  P  S  V  S  I  I  S  D  V  V  P  A  E
TGGGATGATTCAGATGCAAACGTCTGGGGTGAGGACTAGTCAGATGATTGATCGACCCTT   9660
 W  D  D  S  D  A  N  V  W  G  E  D
CTACCGATTGGTGATCGCTATCCTTGCTCTCTGAGAAATAGGTGAGGCGAAACAAAAAAT   9720

AATTTGCATGATAAAAAGTCTGATTTTATGATCGCTATCCTCGCTCTCTGAGAAAGAAGC   9780

GAAACAAAGGCGACTCCTGGACTCGAATCTATAAGATAACAAAGGCGACTCCTGGGACTC   9840

GAATCTATAAGATAACAAAGGCAATTCCAAGACTTGAATCTATAAAAAATTTAGTTAAGA   9900

ATGATTAACGTCCGATCCTAATTCGAATCGAGGCATCTTACCACTCCATTGATAATTATA   9960

TAAGTCAATAAGTCATATAAWAGTATTAAAAACTAAATTGACTTGATCGGTCTATCAAAA  10020

ATMAGATMAAATTGTGTTCATATGTAACATTTTTGTTGTCACAATTAGCTTAATTACATC  10080

TTTCATGTGCAATAACAAAGAAATGATAGGAATTTAGAGATTCCAATTTTTTTGTTGCCA  10140

CAATTAACTTAATTACATCTTTCATTTGCAATAACAAAGAAATGATAGGAATTTAGAGAT  10200

CCAGTGTCAATACACAACCTAGGCCAACATCGAAAGCATAACTGTAAACTCATGCATGAA  10260

GAAATCAGTCGTAAAAATGAATAAATGCGACATAAAAACAAATTGCATGTATCATTAATG  10320

TGACTTAACTACAAGTAAAAATAAATTTAACAAATGTAACTTAACTACAAGTAAAAATAA  10380

ATTGCTTCTATCATTAACAAACAAACAGAATTAAAAAGAAAAAAACATACTAAATCTTAC  10440

CGTCATTCGATAAAAAAAAATACCAAATTCATAATGCAAGGAAAACGAAACGCGTCCTGA  10500
```

FIG. 12 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890
TCGGGTATCAACGATGAAATGGACCAGTTGGATCGACTGCCTGCACAACGTTAGGTATGC  10560
CAAAAAAAAGAACACGATCCTTTGCACCCGTTCGATGATTATCAGTATGTTCACAAAAAA  10620
AACTTAAGTTCATCCCAGTGTACAACAGCCCCAACATCTGCCCCAAGTAACAAAAAACAA  10680
CCAATTTATCTTATTCTTATCTGCCACAAAATAATCGGTTTCACACTATTCTCTTGTTAT  10740
ACAAAATTGACAAGTAGGAAGGAGAGGAGTCATCCAAATAAACGGTGCACGTTCTTTGAG  10800
AAAAGTCTTATTTTTCGTAAGATCCAATTTCAACAAACTTTTCTTCAAGTCAAAATTCCT  10860
GATAGTGTATCTCCTCTCGACGACCTCTTGCATTGAACGATCTCCGCTTATCATGAAAAG  10920
TTGCTTGGATAACAAGTATTGCAAGGGGGGACAGTAGCTATTAAGTTAGTCGGCCCAAG  10980
GAAATGGAGGAGTGATAGTCTCGAATATTATTCACCTCTTTAGCATTACCCGGTCTGGCT  11040
TTAAGGAGTTACGTCTTTTACGCTCGCCAATTTCTTTTTTTAGAATGGTTGGTGTCAAAA  11100
TCGCGAGTTGTGGAAGGTTCAAGTTACTCGATTCGTGATTTTCAAGTATGAGTGGTGAGA  11160
GAGATTCGATATTTTCACGAGGTGTATTCGAGGTCTAGTAGAACGAAGGGTGTCACTAAT  11220
GAAAGTTTCAAGAGTTCATCATCATCTTCTTCTAGTAGATTTTCGCTTTCAAATGAGTAT  11280
GAAAATTCTTCCTCTTTTCTATTGATTTTCTTCATTGTTTTCTTCATTGTTGTGGTTGTT  11340
ATTGAAAAGAAAGAAAATTTATAACAGAAAAAGATGTCAAAAAAAAGGTAAAATGAAAGA  11400
GTATCATATACTTAAAGAGTTGCGTAGAGATAAGTCAAAAGAAACAGAATTATAGTAATT  11460
TCAGCTAAGTTAGAATTC                                            11478
```

FIG. 12 CONTINUED

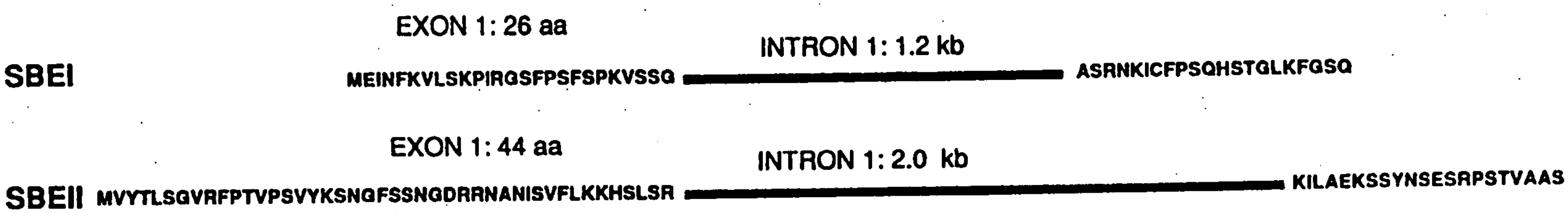

EXON 1: 26 aa
INTRON 1: 1.2 kb
SBEI
MEINFKVLSKPIRGSFPSFSPKVSSG
ASRNKICFPSQHSTGLKFGSQ
EXON 1: 44 aa
INTRON 1: 2.0 kb
SBEII
MVYTLSGVRFPTVPSVYKSNGFSSNGDRRNANISVFLKKHSLSR
KILAEKSSYNSESRPSTVAAS

FIG. 13

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890


GTATACACTCTCTGGAGTTCGTTTTCCTACTGTTCCATCAGTGTACAAATCTAATGGATT     60
 Y  T  L  S  G  V  R  F  P  T  V  P  S  V  Y  K  S  N  G  F

                                    SspI
                                 BsmI
CAGCAGTAATGGTGATCGGAGGAATGCTAATATTTCTGTATTCTTGAAAAAACACTCTCT     120
 S  S  N  G  D  R  R  N  A  N  I  S  V  F  L  K  K  H  S  L

    BsaAI
TTCACgtatgtctcactgtgtttgtggctgtgtgtgtttttttctctgtctttttgtgtt     180
 S  R

                Bsp1286I
                BanII
ttgtgtaattggggctctttaaagttggtattgtgtataccccttttgagtatagtctttg     240

aggaagcaaaatgatgaatcttgattgacattagtaagggttgtaactttttgaagtttg     300

gttaggtgtaattgagtttggcttgtgtgtctgtgtgtcgaggttatttttttggtttgt     360

gttattggggatcttaaaagttggtattgtgtataccccttttgagtatagtctttgagga    420

agcaaaaatgatgaatcttgattggcattagtaaaggttgtagctttttgaagtgtggtt     480

aggtgtaattgagtttggcttgtgtgtctgtgtgttttggaatcctgatgtgtgtcaagt     540
```

FIG. 14

```
        10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890

cctgatatgggtcgaggttctttctttggtttgtgtaattgggggttcttaaaagttggt  600

                                         ClaI
                                         BspDI
attatgtaccttttttaagaatagtgtctgagaaagcaaaatcgatgaattttgattgaca  660

gcatattctttgagaaagcaaaaaatggtgagttttcatggagaaacttgattgacatta  720

ctaaaggtagcaactttttcaactcctgatatgggtcaaggttctttgtttggtttgtgt  780

aatttggggttctttgaagttttgagaaagaaaaattatgatttttcatggagaaatttg  840

                                          PvuII
       AseI                               NspBII
atttacattaataaaggtagtagctttttaaagtgtggtcagctgtaatgagttcagctt  900

           Bsp1286I
           BanII
           ApaI   NdeI
ggtttaaaggggcccctacatatggtgctttctggtgagatatttgttgctccaccatac  960

gagttataagaatcatagtgttaggatctttttcttttttttttcatttttcacttgac  1020

tagctactagaggagtgatcttgacggcggaaaatcttagaaaggggaaggttgtttgca  1080
```

FIG. 14 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890

                  Esp3I                         BsaBI
tcaactggtgttatatgtgcaaggagacgggagatgatgtagatcatcttcttcttcatt   1140

gtggtctttccatgaggttatgatgtgatatgtttgaatggtttggtacttcttggctat   1200

                                           EarI
gccaagaactgtgaaagaattgatattcagttggaagtgtggagttggaagagtggaaga   1260

attgacacttggttccattagctttaatgtgggtggtgtggagagagagagaaataggag   1320

                                                       EcoRV
agcttttgagggggtagagttgagctttcctcagttgagaagtagcctttgatatctttt   1380

                        EcoRI  MunI
ttttttttttttgtacacccatagaattcccaattgtatagaagattgggtggagtttgt   1440

agagaatcatcttttgtagtagattctttaccttttggtatatccattgtatacagccag   1500

 StuI
gcctttgactatgtttatgaatgaatatacattacttgaaaaaaaaagaagtgaagccag   1560

tctgttgtacctttgtagacaatgttgttgcagcatcttgataattccctgaaaattgtc   1620
```

FIG. 14 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890

tccctgaaggaatagtttggttgatattgattatttcttggtttgtttaattcggtgttc  1680

ttgaaggccattttaaatcctttgacattgttaaaggtgtttacaagtgttggtctgggt  1740

ttaaaagcacctcttgtatggtgctttctggagtgatctttcttcctccaaaagagaagt  1800

                                      BclI BglII
tgcaagaatcagtgtgtgtactttttttctcttgtatgatcagatctttttttcaattttttc  1860

cgttttagttgatttatccatatagtgaaagttggtgtcatagttgctgtttgtggactt  1920

cctgtaaaagtttttttgatatacttaaaaaattgtcacacagaagaaagagtttttttacc  1980

   AflII
attacttaagctagatgggactgtttgattcttagaccaaataatgaacctttttgttct  2040

   AflIII
cttaacgtgtacttgaaatagtttggtaaaattgtgataggaaaaaagataattcttgat  2100

                                              EarI
tgcttttggagcatcacttctaatcataaaagtctttgctctcttcaaccatgaatgata  2160
```

FIG. 14 CONTINUED

```
         10        20        30        40        50        60
123456789012345678901234567890123456789012345678901234567890

aattggacacttatgtggccctaagttgctctcagtagtggtctttaattgtggagatat    2220

                                    BglII     BbsI
aactaatctgatatatgtatgtagGGAAGATCTTGGCTGAAAAGTCTTCTTACAATTCCG    2280
                         K  I  L  A  E  K  S  S  Y  N  S  E

            SfcI
AATCCCGACCTTCTACAGTTGCAGCATCG                                   2309
  S  R  P  S  T  V  A  A  S
```

FIG. 14 CONTINUED

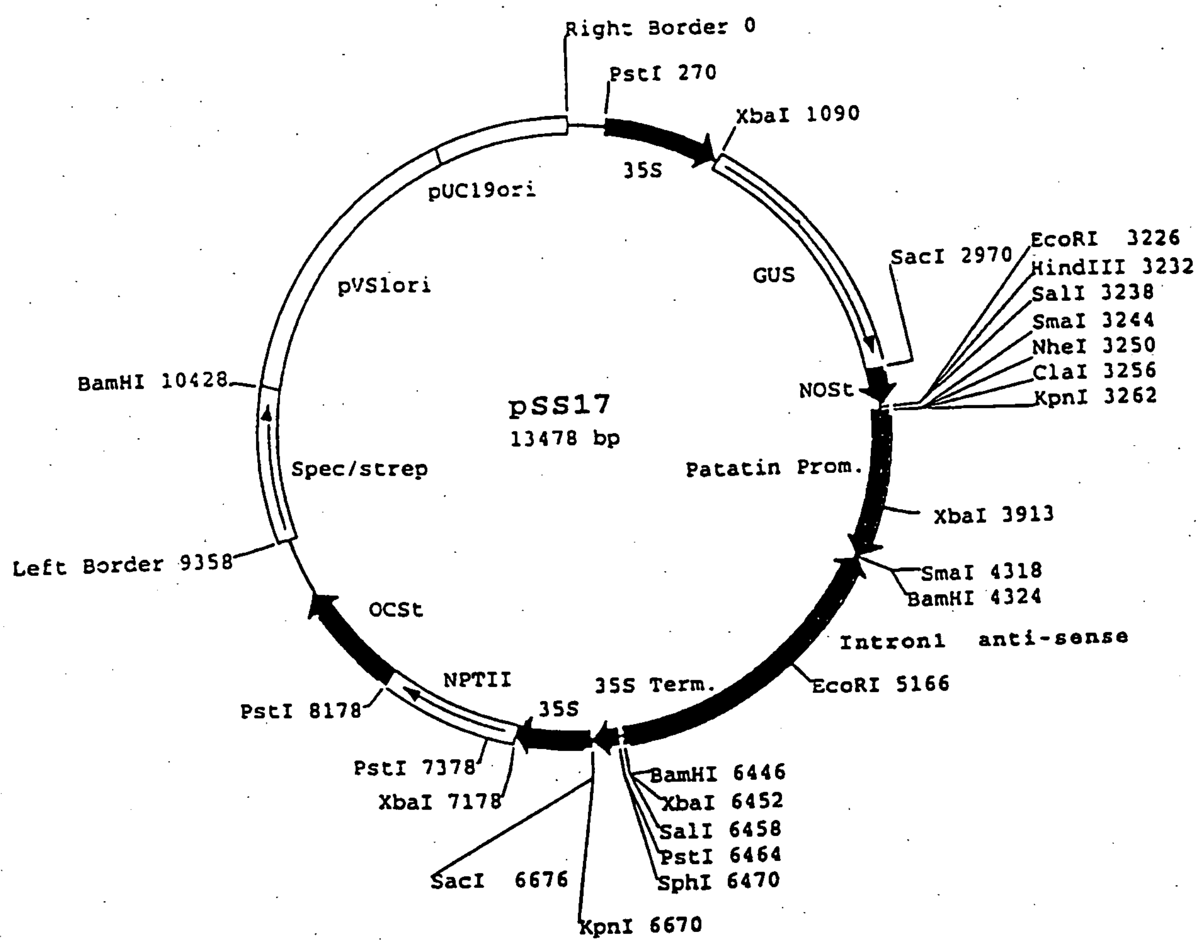

Right Border 0
PstI 270
XbaI 1090
35S
GUS
SacI 2970
EcoRI 3226
HindIII 3232
SalI 3238
SmaI 3244
NheI 3250
ClaI 3256
KpnI 3262
NOSt
Patatin Prom.
XbaI 3913
SmaI 4318
BamHI 4324
Intron1 anti-sense
EcoRI 5166
35S Term.
BamHI 6446
XbaI 6452
SalI 6458
PstI 6464
SphI 6470
KpnI 6670
SacI 6676
35S
XbaI 7178
PstI 7378
NPTII
PstI 8178
OCSt
Left Border 9358
Spec/strep
BamHI 10428
pVS1ori
pUC19ori
pSS17
13478 bp

FIG. 15

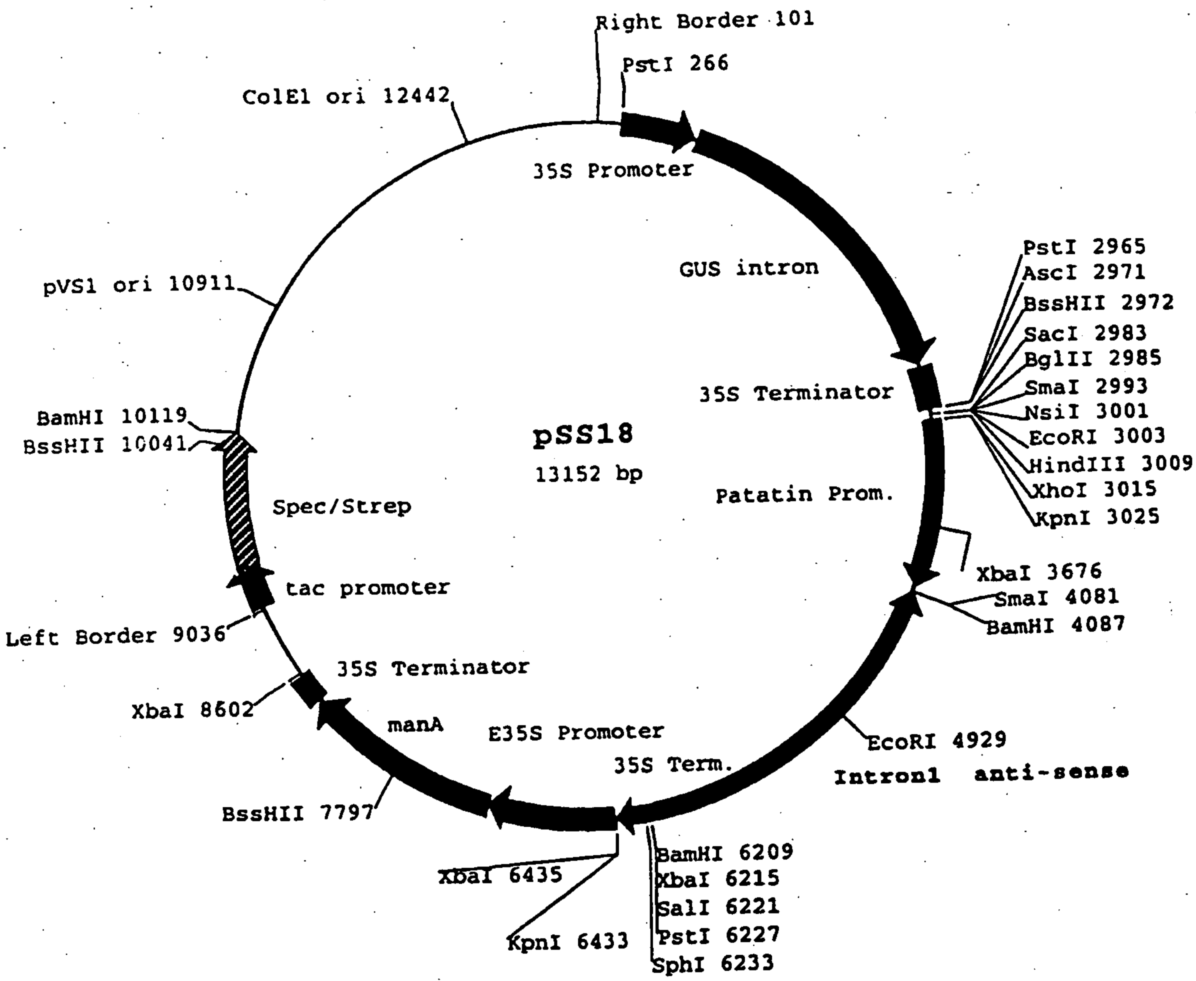

Right Border 101
PstI 266
ColE1 ori 12442
35S Promoter
GUS intron
PstI 2965
AscI 2971
BssHII 2972
SacI 2983
BglII 2985
SmaI 2993
NsiI 3001
EcoRI 3003
HindIII 3009
XhoI 3015
KpnI 3025
35S Terminator
pVS1 ori 10911
BamHI 10119
BssHII 10041
pSS18
13152 bp
Patatin Prom.
Spec/Strep
XbaI 3676
SmaI 4081
BamHI 4087
tac promoter
Left Border 9036
35S Terminator
XbaI 8602
manA
E35S Promoter
EcoRI 4929
35S Term.
Intron1 anti-sense
BssHII 7797
BamHI 6209
XbaI 6215
SalI 6221
PstI 6227
SphI 6233
XbaI 6435
KpnI 6433

FIG. 16

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/00270

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N9/10 C12N15/11 C08B30/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C08B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 04112 A (DANISCO ;POULSEN PETER (DK)) 6 February 1997 cited in the application see the whole document ---	1-21
X	WO 97 04113 A (DANISCO ;POULSEN PETER (DK)) 6 February 1997 cited in the application see the whole document ---	1-21
Y	WO 96 34968 A (NAT STARCH CHEM INVEST ;COOKE DAVID (GB); DEBET MARTINE (GB); GIDL) 7 November 1996 cited in the application see page 5, paragraph 3 - paragraph 4 see page 9, paragraph 2 - page 10, paragraph 1	1-21
X	see page 11, paragraph 3 ---	17-19
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 May 1998	09/06/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Chakravarty, A

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte   nal Application No
PCT/IB 98/00270

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WO 92 11375 A (AMYLOGENE HB) 9 July 1992 cited in the application see the whole document	1-21
Y	WO 94 09144 A (ZENECA LTD) 28 April 1994 see page 10, line 1 - line 18	1-21
Y	WO 92 15680 A (UNIV TEXAS) 17 September 1992 see page 6, line 17 - line 28	1-21
X	EP 0 240 208 A (CALGENE INC) 7 October 1987 see page 3, line 10 - line 13	15

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte: ɔnal Application No
PCT/IB 98/00270

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9704112 A	06-02-1997	AU 6614596 A	18-02-1997
		EP 0839202 A	06-05-1998
WO 9704113 A	06-02-1997	AU 6614696 A	18-02-1997
		EP 0839203 A	06-05-1998
WO 9634968 A	07-11-1996	AU 5509996 A	21-11-1996
		EP 0826061 A	04-03-1998
WO 9211375 A	09-07-1992	SE 467160 B	01-06-1992
		AU 9109791 A	22-07-1992
		EP 0563201 A	06-10-1993
		PL 169859 B	30-09-1996
		SE 9004095 A	01-06-1992
WO 9409144 A	28-04-1994	CA 2146998 A	28-04-1994
		AU 690517 B	30-04-1998
		AU 2696492 A	09-05-1994
		EP 0664835 A	02-08-1995
WO 9215680 A	17-09-1992	AU 663702 B	19-10-1995
		AU 1570492 A	06-10-1992
		CA 2108144 A	07-09-1992
		EP 0575518 A	29-12-1993
		US 5747469 A	05-05-1998
EP 0240208 A	07-10-1987	AT 114168 T	15-12-1994
		AU 1301792 A	03-09-1992
		AU 618234 B	19-12-1991
		AU 7059787 A	01-10-1987
		DE 3750755 D	22-12-1994
		DE 3750755 T	18-05-1995
		EP 0458367 A	27-11-1991
		ES 2066759 T	16-03-1995
		JP 2702921 B	26-01-1998
		JP 62296880 A	24-12-1987
		JP 10052283 A	24-02-1998
		US 5107065 A	21-04-1992
		US 5453566 A	26-09-1995
		US 4801540 A	31-01-1989